AD-A286 432

Bioacoustic Signal Classification in Cat Auditory Cortex

Final Report for Office of Naval Research Grant N00014-91-J-1317

Principal Investigator: Christoph E. Schreiner, Ph.D.,MD

Associate Professor in Residence

Department of Otolaryngology, Coleman Laboratory,

W.M. Keck Center for Integrative Neuroscience

University of California, San Francisco, CA 94143-0732

Tel.: (415) 476-2591, Fax: (415) 476-1941, e-mail: chris@phy.uscf.edu

DTIC
ELECTE
NOV 2 3 1994
F

TABLE OF CONTENT:

94-35593

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

General Rational & Significance:

The classification of signals from the surrounding world is one of the most important, complex, and, at the same time, least understood capabilities of the mammalian brain. Traditionally, attempts to comprehend categorizational mechanisms have utilized the physical depiction of the incoming signal, e.g., the spectrogram of a sound or its waveform, as the starting point for an analysis. Extensive efforts to unravel 'invariant features' in speech, the most complexly categorizable signal, have largely failed to disclose reliable and valid common features in the physical representation of the signal that can account for the remarkable robustness of the human speech categorizer in the presence of speaker variability, noise interference, and partial signal omissions.

An alternative approach to elucidate fundamental principles underlying signal classification utilizes the biologically preprocessed representation of signals in the mammalian brain. This approach makes use of biologically realized and physiologically definable processing strategies and resultant signal representations in the CNS that undoubtedly contribute to the robustness of the signal categorizer/classifier in mammals.

A necessary condition to achieve this goal is the existence of and the access to a complete and systematic biological representation of the input continuum. So far, only two stages along the auditory pathway have been explored with enough detail to provide such a biological representation of the input: the auditory nerve and the primary auditory cortex (AI). Whereas there exist a multitude of isolated observations in the intervening stations, only in the two before mentioned levels have there been attempts to reconstruct a *complete distributed ensemble representation* for test stimuli. Although the systematic analysis of processing properties is much more advanced in the relatively uniformly organized auditory nerve, the more complex signal representation in the auditory cortex provides a much more suitable level for elucidating the foundations of biological signal classification. The primary reason for this assessment is that more complex processing strategies, e.g. noise reduction, level tolerance, and feature enhancement, have to be accomplished mainly beyond the level of the auditory nerve. In addition, recent studies of the cat primary auditory cortex in our laboratory and elsewhere clearly indicate that the form of signal representation in this cortical field is that of a generalized and possibly uncategorical representation not unlike that of the primary visual cortex and in contrast to a specialized, feature combination sensitive, categorical representation that has been shown for a number of subfields surrounding the bat primary auditory cortex.

The signal encoding and representation in the primary auditory cortex of cats and owl monkeys, the only species that have been studied with sufficient detail besides the bat, shows a number of systematic and spatially distributed features that may provide robustly enhanced properties that form the basis for a tolerant signal classification. Other auditory cortical fields, surrounding AI, appear to have slightly (AAF) or grossly (AII) different response characteristics from those seen in AI. Biologically, only parallel and cooperative processing may lead to the desired signal analysis and categorization. For our analytical attempts, it is similarly likely that only a combination of information from all cortical fields may provide us with all necessary cues for a n animals strategies in signal categorization. Nevertheless, it is hypothesized that the representation of acoustical signals in the primary auditory cortex is an ideal starting point to explore the cues available for signal categorization after extensive biological 'pre'-processing. Based on a systematic cortical map of signals, the strategies that lead to such a representation as well as their consequences for selection of classification algorithms for speech signals and sonar signals can be explored.

These studies attempted to examine the representation of acoustical signals in the primary auditory cortex of mammals and to derive detection and classification strategies based on the biological signal representation of complex signals in the CNS. We postulated that the subcortical/cortical processing of signals results in a systematic and robust signal representation in the primary auditory cortex that is more suitable for signal detection and classification schemes than acoustical/spectral signal representations. Models of central nervous representations were to

A 237646

A 1

be developed and applied to the transformation and classification of complex signals. The general objectives for the research conducted in the last funding period were as follows:

Objective 1: To explore the physiological principles underlying the spatio-temporal representation of simple and complex signals in the primary auditory cortex;
Objective 2: To create a formalized model/simulation of the cortical representation of complex signals based on fundamental principles of spatio-temporal processing;
Objective 3: To utilize the formalized cortical representation as the input (front end) to a self-organizing signal classifier and as training pattern for the output of a dynamic neural network.

In the following section, the progress toward these objectives over the last funding period will be briefly reviewed.

Objective 1: In order to explore in more detail the physiological principles underlying the spatio-temporal representation of simple and complex signals in the primary auditory cortex we completed a series of experiments that utilized narrow- and broad-band signals to derive local filter properties of single cortical neurons and small groups of cortical neurons. The underlying assumption is that the characterization of local filter properties is necessary and sufficient to predict the cortical neuronal response and spatial response distribution for arbitrary acoustical signals.

Narrow-band stimuli and frequency modulation (in collaboration with Julie Mendelson and Mitchel Sutter): The spatial distribution of several response parameters to pure tone stimulation (threshold, best level, dynamic range, monotonicity; Schreiner et al., 1992) and frequency sweeps was evaluated (Mendelson et al., 1993). In addition, we obtained and compared the topographical distribution of several aspect of pure-tone receptive fields, namely best frequency, excitatory bandwidth, response threshold, and monotonicity of rate/level functions (Schreiner and Sutter, 1992, see Appendix; Sutter and Schreiner (in press) see Appendix) for single and multiple unit recordings in the same animals. The purpose of this comparison was to guide the interpretation of multiple unit mapping experiments by relating spatially integrated response measures to the properties of their generators. Among the main findings of these studies are:
a) All studied parameter show a nonuniform spatial distribution along the isofrequency axis of cat auditory cortex.
b) Several of these parameters, especially FM direction and FM speed were correlated with the spatial distribution of the excitatory bandwidth suggestive of a close relationship between some dynamic receptive field aspects and static receptive field properties. Other parameters, e.g. threshold and excitatory bandwidth, appear to be independently organized.
c) In dorsal AI, single unit and multiple unit properties were similar with regard to the bandwidth measures but different for the intensity parameters. In ventral AI, both measures showed similar results for intensity coding parameters but differences between the multiple and single unit bandwidth organizations. These findings provide further evidence that the dorsal and ventral half of AI process information differently. Differences between single unit mapping and multiple unit mapping could be accounted for by differences in the local CF scatter across AI and differences in the local scatter of response threshold. In addition, the maximum response strength of single neurons to pure tones varies in AI with nonmonotonic neurons usually contributing more spikes, at best level, than monotonic neurons.
d) One region in the center of the dorsal-ventral extent of cat AI appears to have a constellation of parameters that may subserve a special signal detection purpose. Near the border between ventral and dorsal AI, neuronal recordings are sharply tuned for frequency (high Q values) and for amplitude (high degree of nonmonotonicity). In addition, the frequency scatter and threshold scatter is minimal and the neurons show the lowest response thresholds. This constellation fulfills the conditions to serve as a region for fine spectral/amplitude filtering which will only respond to frequency-banded components with intensities just above the animals threshold. Since

3

it has been shown that addition of background noise will shift the rate/level function of nonmonotonic neurons to higher levels, this region maintains its ability to detect a signal close to threshold even in the presence of background noise (see also below).
e) A study of the spatial organization of the primary auditory cortex in owl monkeys (Recanzone and Schreiner, in preparation) revealed similar principles as seen in the cat, i.e. bandwidth, threshold, and latency gradients along the isofrequency axis as well as well expressed, elongated binaural interaction patches (or bands). These findings suggest, that the organization of AI seen in the cat reflects general organizational principles including that of human auditory cortex.

Broad-band stimuli (in collaboration with Ms. Barbara Calhoun and Diane Keeling): Narrow-band stimuli are a versatile tool to easily characterize neuronal receptive fields, since the stimulus is chosen to be clearly definable regarding temporal, frequency, and intensity characteristics. Resulting temporal, frequency and intensity transfer function for a receptive field usually keep two of these aspects constant and vary the third one. Naturally occurring stimuli, however, usually consist of a wide range of temporal, frequency and intensity combinations that render it extremely difficult, if not impossible, to predict cortical neuronal responses based on the narrow-band transfer functions accumulated in frequency response areas. For example, the strong level dependence of the shape of a frequency transfer function (a cross section through the frequency response area) is principally different from a simple FIR filter that can be applied to a complex signal at any level. To what degree can frequency response areas indeed account for the response of complex, broadband signals? In a first attempt to address this question, we developed a stimulus that ultimately should enable us to provide an answer. The chosen stimulus consist of a broadband signal, either noise or a harmonic series, whose spectral envelope is sinusoidally modulated (on a logarithmic frequency scale to make the envelope features largely frequency shift invariant). Parameters of the spectral envelope are 'frequency': ripple density measured in 1/octave; 'amplitude': modulation depth measured in dB; and 'phase': frequency position of maxima or minima of the spectral envelope, measured in degree. These 'ripple' spectra are essentially the same as sinusoidal luminance gratings used for the investigation of the visual system. Their relevance for the auditory system is threefold. First, they allow to study directly, systematically, and parametrically properties of cortical cells in response to a broadband, complex signal. Second, their form resembles closely an important aspects of all vocalizations, namely vocal tract resonances that result in spectral peaks or formants in human speech as well as in animal vocalizations. Third, the sinusoidal nature of the spectral envelope allows, with some caveats, a system theoretical approach that can provide an estimate of the frequency response area based on a broadband stimulus (for details see below).

By studying responses to different spectral envelope frequencies, we obtained information regarding the processing of elementary signals that may lead to the understanding of the processing of arbitrary spectral envelope waveforms, including formants in vocalizations, and 'spectral edges' such as notches due to resonances in the outer ear. The spectra with sinusoidal envelopes or 'ripple spectra' applied in this study have the following characteristics: the carrier consists of a harmonic series (f_0 ranging from 50 to 200 Hz) with a 6dB/octave decline of the component amplitudes (120 to 255 components); the bandwidth of the stimulus is usually 3 octaves; the spectral envelope of the signal is represented by a sinusoid on a logarithmically scaled frequency axis (since the basilar membrane essentially provides a receptor surface with a logarithmic frequency distribution); the frequency of the envelope sinusoid is referred to as ripple density (ripples/octaves); the modulation depth of the envelope (ripple depth) is linear on a dB scale.

In this study, the geometrical center of the band-limited signal was always at a maximum of the sinusoidal spectral envelope. The center was positioned at the CF of each cortical neuron and the ripple density or the frequency distance between spectral peaks was systematically varied. The resulting 'ripple transfer function' was reconstructed for different modulation depth and overall intensities. For the majority of neurons, the ripple transfer function was a bandpass and a 'best' ripple density could be defined. The remaining transfer functions appeared to have a

4

lowpass characteristic, at least for those ripple densities employed in this study (0.3 to 8 ripples/octave). Best ripple densities ranged from 0.3 to 4 ripple/octave with a mean around 1 ripple/octave.

The spacing of formants in human speech can be expressed in ripple density by calculating the reciprocal of the spectral distance between two formants measured in octaves [e.g. 1/ (2 octaves) corresponds to a ripple density of 0.5]. Most formant spacings in human vowels are in the range of 0.6 to 1 ripple/octave with the majority below 3 ripples/ octave. Examination of the ripple densities present in the vocalization of cats also shows a prominent peak around a ripple density of 1 ripple/octave.

These findings suggests that neurons in the auditory cortex can serve as filters for different spatial frequencies or spectral envelope frequencies and that the filters are predominantly tuned to spectral envelope features and frequencies that are prominent in their behaviorally relevant acoustic environment.

Spatial mapping of responses to ripple spectra along the isofrequency domain of AI revealed a systematic shift of the best ripple density in multiple unit responses from central AI to dorsal AI. The shift paralleled the previously described variation of integrated excitatory bandwidth with sharply tuned locations and high best ripple densities near the dorso-ventral center and broader tuning curves with lower best ripple densities toward the dorsal end of AI. Ventral AI showed a less systematic distribution of ripple densities.

The findings indicate that the auditory cortex is well suited for the analysis of complex, broad-band stimuli. The results support the hypothesis of a systematic 'spatial frequency ' or 'spectral envelope frequency' representation in AI. This representation is oriented orthogonal to the frequency axis and, therefore, is available for every frequency channel. The representation is covering the main spectral envelope frequency range found in complex signals such as speech and can serve as an analyzer and classifier for arbitrary spectral waveforms, including important classes such as formants and spectral notches. The topographic distribution of different spectral envelope frequencies across AI supports the psychophysical finding that separate independent 'ripple frequency channels' exist that can be adapted (long duration of stimulation with a given ripple density will degrade the discrimination of that ripple frequency channel but not others; Hillier, 1992)).

Previously, we have shown that the cortical organization of frequency representation can be altered through experience (see Recanzone et al., 1993). In an attempt to demonstrate that the cortical representation of information within the isofrequency domain can be adjusted through experience as well, we trained several cats to discriminate between two stimuli that possessed the same spectral envelope frequency but with different locations of the spectral maxima (Keeling et al. 1995a (submitted) see Appendix). It was hypothesized that exposure to a specific spectral envelope frequency that is made behavioral relevant to the animal would emphasize and refine the cortical representation of that envelope characteristic. It was found that some but not all animals were capable of improving there ability to discriminate those sounds over the period of several month of training. Electrophysiologically, it was found that the ripple transfer functions of neurons in AI of trained cats were indeed different from those in untrained control animals (Keeling et al. 1995b (submitted) see Appendix). The peak of the compound ripple transfer functions shifted toward the trained envelope frequency and, in addition, an increased responsiveness to higher spectral envelope frequencies (broadening of ripple transfer functions) was observed.

These findings suggest, that behavioral training (not just exposure) to specific aspects of complex sounds can result in an improvement of the psychophysical capacity of a subject (time course: at least four to eight weeks of training) and that cortical neurons can alter their response characteristics as a consequence of that training. It is likely, that the improvement in the cortical representation of the trained signal features is a necessary condition for improvement of the psychophysical capacity. However, at this time it is not clear whether the observed cortical changes are also sufficient for psychophysical improvement to occur.

5

Temporal and Binaural Response Properties (in collaboration with Mitchell Sutter and Julie Mendelson): The neuronal responses to tones, broadband transient stimuli, and frequency-modulated (FM) sweeps were mapped in the primary auditory cortex (AI) of barbiturate-anaesthetized cats (Schreiner et al., 1995 (submitted) see Appendix). The spatial distribution of the final two response parameters in this series of experiments, onset latency and binaural response, were determined in four cases. The functional relationship between these two parameters and the previously reported spectral (Q_{10dB}, Q_{40dB}, broadband transient stimulus), intensity (threshold, strongest response level (SRL), dynamic range and monotonicity) and temporal (FM sweep direction and speed) parameters was subjected to both a global and regional analysis. Onset latency responses were systematically distributed along the dorsoventral/isofrequency axis of AI such that units with shorter latencies were located in the central region while units with longer latencies were more often found in the dorsal and ventral portions of the cortex. As shown by other investigators, alternating bands or patches of EE and EI units were also distributed across the cortex. The results of a point-by-point analysis revealed a lack of dependency of onset latency on binaural response type. However, when the relationship between onset latency response and the remaining 10 response parameters was examined, the correlations with the following parameters were found to be statistically significant: Q_{10dB}, Q_{40dB}, broadband transient response, strongest response level (SRL), monotonicity, and preferred FM sweep direction. These correlations suggest that units that respond with shorter onset latencies are more sharply tuned, do not respond well to broadband stimuli, have a higher degree of nonmonotonicity, and prefer FM sweeps that change from a high to a low frequency. Binaural response type was found to be significantly correlated with Q_{40dB}, broadband transient response, threshold, SRL, dynamic range, and preferred FM sweep speed. These correlations imply that EI neurons tend to have sharper tuning, weaker responses to broadband stimuli, lower thresholds, lower SRLs, a narrower dynamic range, higher degree of nonmonotonicity, and prefer slower FM sweeps than EE units. The systematic distribution and the functional relationship between these response parameters may provide the representational basis for the detection and identification of specific features in the animal's natural environment. The presence of a number of response parameters that appear to be systematically organized within AI may provide a set of criteria by which the dorsoventral extent of AI can be delineated. With the exception of binaural response, changes in the distribution of other response features, e.g., onset latency, are gradual on a global scale, but can be quite variable on a local scale. However, they may not necessarily provide an unequivocal set of criteria that would allow for a clear demarcation of the borders to areas neighbouring dorsal and ventral AI.

On a functional level, the systematic distribution of these (and perhaps other as yet unknown) response parameters may reflect the representational basis for the detection and identification of specific features of sounds arising from the natural environment. Schreiner and Sutter (1992) have recently suggested that the region in AId (dorsal to the Q_{10dB} peak) may be better suited for integrative analysis of broadband stimuli by responding in a relatively undifferentiated manner to tones or spectral peaks of different frequencies. This is supported by the observation that cells in this subregion tend to be more broadly tuned and to have a high responsiveness to broadband stimuli such as clicks. In contrast, AIv may be better suited for differential analysis of the spectral properties of broadband stimuli since units there are characterized by sharp tuning which is distributed over varying ranges of CFs and which do not seem to respond to broadband stimuli as well as cells in dorsal AI.

Receptive field characterization (in collaboration with Dr. Xioaquin Wang): One of the main problems encountered in constructing filters to model the function of cortical neurons is that it is extremely difficult to predict the response of neurons to complex signals because the exact timing of an action potential occurrence depends on a number of parameters that have not been well described, such as the frequency dependent latency profile in the excitatory portion

6

of the tuning curve or the influence of inhibitory influences on the arrival time of spikes. Accordingly, it is impossible to precisely assign spikes in the response to a complex signal to a certain event in the stimulus. The technique of reverse correlation, however, provides a means to obtain a probability measure of which signal portion produced a given spike.

The technique we developed is similar to the reverse-correlation method used by Jones and Palmer (1987) in the visual system and by Eggermont et al. (1983) in the auditory periphery. However, the approach had to be somewhat modified. In the peripheral auditory system, the reverse correlation method averages the signal waveform preceding every neuronal event during a long period of continuous stimulation with either complex signals or noise amplitude modulated with a lowpass noise. The result is a pre-event stimulus ensemble (PESE) reflecting the probability distribution in the stimulus that precedes a neuronal event (spike). In the periphery, the phaselocking of the response to the stimulus waveform is a crucial for successfully deriving the PESE. The timing of cortical spikes, however, is too imprecise and does not phaselock to the carrier component of stimuli. A direct reverse correlation approach to the raw stimulus waveform would result in a flat PESE due to the temporal jitter of the spike occurrence. By contrast, if the reverse correlation is done on the basis of the spectral and temporal envelopes of the complex (natural) sounds, it will be successful in characterizing the spectro-temporal receptive field of neurons since the time jitter in cortical responses is small compared to the slow temporal events the cortex is able to follow (e.g. Schreiner an Urbas, 1988).

Figures 1 and 2 give two examples of the result of this modified reverse correlation method. A narrowly tuned (frequency response area in Fig 1D) and a more broadly tuned neuron (Fig. 2D) were studied. The spike arrival times during stimulation with pure tones and a set of natural vocalizations were obtained. The PESE was reconstructed by first constructing a 128-channel envelope/time representation of the signals, i.e. each signal was split into 128 narrow frequency bands and the temporal envelope for each band was obtained by lowpass filtering each band (cutoff: 200Hz, well above the normal cutoff frequency for cortical neurons). Then, the signal content of each band was accumulated for a 30 msec period immediately prior to the arrival of each spike. Since the spectral content of the vocalization did not cover the whole frequency range of 24 kHz used in these examples with the same probability, we subtracted from this raw PESE the PESE that would emerge if the same number of spikes elicited by the stimulus would occur randomly during the stimulus. This results in the spectrally corrected PESE. Figure 1A and 2A show the PESE for the vocalizations. Note that the red areas correspond to high probability values that a signal of that nature occurred before the spike, green corresponds to the signal occurrence probability without causal relationship between signal and spike, and blue corresponds to a reduced probability, i.e. during those sections, a signal had to be absent in order for a spike to occur (presumably reflecting inhibitory properties of the receptive field). In Figure 1A, the region of elevated probability is located between 15 and 20 ms before the spike occurrence and is restricted to a fairly narrow frequency range. This range corresponds well to the timing and spectral properties of the excitatory pure tone response (see Fig. 1B and 1D). However, other features of the receptive field emerge that are not seen in the pure tone response, among them are the blue regions of inhibition surrounding the excitatory region on both sides, and a slant in the probability distribution indicating that certain frequency modulation directions and speeds may be optimal for this cell.

Objective 2: One of the objectives of our work on the auditory cortex is to be able to predict the response to complex signals from the receptive field information. In order to accomplish that, we have to develop a formalized version of a filter model based on the experimentally obtained receptive field properties and spatial distributions. The applications of such a formalized model are manifold. Of special interest to us is their potential use as a front end for a biological based signal classifier, their use as a trainer for network models, and their ability to predict spatial distribution

7

pattern of activity to complex signals which is necessary, for example, in the interpretation of optical recordings from the auditory cortex or of output patterns of selforganizing networks.

We developed a filter model that was based on the experimentally obtained characteristics of 133 cortical neurons distributed over a narrow frequency range (center frequency 5 kHz, range approximately 0.5 octaves) and sampled along the entire extent of the dorsal-ventral axis of AI. A network of 133X21 filters was constructed that simulated an area of approximately 6X4 mm of AI. The distribution of the activity of the model neurons was determined after the application of tone bursts at CF for 20 different sound pressure levels (50 to 100dB).

The receptive fields of the model neurons were realized as a set of 20 level dependent bandpasses. The width and the shape pf the passbands of the neurons were first reconstructed from experimentally determined frequency response areas by using their response threshold, best level, rate/level monotonicity, Q-10dB, and Q-40dB information. The lower and upper limits of the passband at different sound pressure levels were calculated from CF, threshold, Q-10dB and Q-40dB, under the assumption that the frequency response area are symmetrical around CF. Taking the logarithmic scaling of the frequency axis into account, the respective upper (f_u) and lower (f_l) cutoff frequency values were calculated for 10dB and 40dB above threshold:

$$f_l = CF * (\sqrt{1 + 0.25 * Q2} - \{1 / (2 * Q)\}) \text{ for } Q = Q\text{-}10dB \text{ or } Q\text{-}40dB$$
and
$$f_u = CF^2/fl.$$

Thereafter, the frequencies were transferred into logarithmic coordinates and straight lines were drawn between the minimum threshold value at CF to the f_u and f_l for Q-10dB and from there to the f_u and f_l for Q-40dB. Cutoff frequencies above 40dB were extrapolated by extending the border of the tuning curve with the same slope as between the Q-10dB and Q-40dB positions. The shape of the passband at a given level is not flat but the magnitude is frequency c pendent. In order to determine its shape, at first the magnitude at CF was calculated for different sound pressure levels. Below threshold the magnitude was set to zero. From threshold to a transition point 20dB above threshold the magnitude increased linearly (dynamic range), and above the transition point the magnitude changed according to the nonmonotonicity value of the neuron. To compare the activities of different neurons, the magnitude at the transition point was set to 100. After determination of the magnitude of the passband at CF, the magnitudes at the other frequencies in the passband was determined for all sound pressure levels by drawing a line from the lower boundary of the passband to the maximum at CF and from this point to the upper boundary of the passband. Since logarithmic frequency coordinates were used, the passband consisted of two parts of exponentially shaped curves.

After the reconstruction of the frequency response area, for each sound pressure level equivalent bandpass filters were designed. To be able to filter different signals with different energies, the amplification of the neuron for each sound sound pressure level was determined by dividing the magnitude of the filter through the amplitude of the sound pressure level. Since the passbands were non-flat we used the modified Yule-Walker method to design recursive digital filters which provide an optimal least square fir to the shape of the filter.

The network was constructed by ordering the properties of the model neurons according to their position along the isofrequency line and setting the CF of the neurons to a fixed value. The center row of the network was set to a CF of 5 kHz. In order to expand the network to different CFs, the isofrequency line was duplicated and the passbands of all neurons in the line was shifted by a specific CF factor to cover a range of neurons with CFs between 1.25 and 20 kHz (+/- 2 octaves).

The network was stimulated with 50ms tone bursts at 21 different frequencies (1.25 to 20kHz in equidistant logarithmic scales) and 20 different amplitudes (5 to 100 dB SPL) were generated. First the power of the signals were determined, and then the signals were bandpass filtered with the appropriate digital filter. The relative activity of each neuron was defined as the

8

maximal amplitude of the filtered signal after cessation of the overshoot to the onset of the input signal, i.e. in the range 5 to 50 ms after toe tone onset.

Figures 3,4 and 5 show the resulting activity distribution of the 133X21 model neurons. On the x-axis, neurons with different frequency response areas are plotted, representing the isofrequency domain. Small numbers correspond to locations near the dorsal end of AI, while high numbers stand for neurons located near the ventral part of AI. The neurons are approximately equidistant along this axis. Along the y-axis, neuron differ in their CF, such that -2 indicates neurons with CFs of 1/4 of the neurons at zero, i.e. they are 2 octaves below the center values. Because of the dimensions of cat AI, the length of the y-axis corresponds to approximately 4 mm. On the z-axis, the relative activity of the model neurons at the different locations is plotted for three different sound pressure levels (10dB, 45dB and 90dB) of tone burst stimuli of 5kHz (corresponding to frequency position zero). Note that at the low sound pressure levels active neurons are chiefly found around the y-axis values of zero, and that more neurons in the center of the x-axis are active. This corresponds to the low threshold region found in the center of AI. At intermediate sound pressure levels (45dB, Figure 4) almost all neurons with CF = 0 are active and the range of active neurons is broader. Within the isofrequency domain, neurons at the ventral and dorsal end of the network are more active than neurons at the center due to threshold and monotonicity properties of the neurons. At high sound pressure levels (90dB, Figure 5) many neurons with different CF at the ventral and dorsal extremes of the network are activated while fewer neurons in the center of the network are active. This spread of activation at high levels corresponds well to the actual behavior of auditory cortex.

Despite the simplicity of this formalized model of cortical neurons, a number of behaviors of the activity distribution reflect the actual findings of cortical activity distributions. However, a number of improvements need to be added to the model in order to allow modeling of signals other than pure tones. There is no need for the introduction of inhibitory sidebands for the prediction of pure tone responses since the pure tone receptive fields include all of the necessary information. By contrast, the generation of responses to broadband signals has to include filter properties that take into account inhibitory interactions.

Objective 3: Two approaches to selforganizing models of cortical activity or parameter distributions have been initiated over the previous funding period.
Connectivity model (in collaboration with Koichi Samishima): The mammalian central auditory system is a complex system with huge numbers of variables contributing to its functions. It is of common sense that with adequate selection of parameters values is it possible to generate any kind of behavior from a system with high degrees-of-freedom. With the current limited computer power resources, it is necessary to introduce some reasonable simplifications into the model. The goal of this work is to reproduce the response characteristics and distribution of primary auditory cortex based on available neuroanatomic connectivity information for the auditory cortex and simple neural network cell model with a Hebbian-like learning rule. It is postulated that, from this starting point, a selforganizing process can recreate some of the features experimentally observed. It is hoped that the outcome of this approach will provide some insights to and predictions of auditory system physiology that might lead to further experimentally addressable questions.

In this modelling we are primarily interested in the understanding the cortical representation of responses to different dimensions of sound, e.g. bandwidth and intensity, across AI. The proposed neural network has five layers, one cochlear excitatory input, two subcortical layers and two cortical layers, as depicted in Figures 6, 7 and 8. The first layer consists of a one dimensional array of 15x1 cells representing the cochlear band-pass filter and hair cells. The subcortex has two layers, one excitatory and one inhibitory, of 15x10 cells. The cortex is represented by the same number of 15x10 excitatory and inhibitory cells. Each cell of the cochlear layer projects to a band of 3 by 10 subcortical excitatory (SE) cells; as shown in Figures 6 and 7 preserving the isofrequency representation along one dimension of subcortical layer. Each SE cells sends efferents to excitatory cells (CE) of the cortical layer with the isofrequency band oriented ventro-dorsally (Figure 6), with increasing spreading of connection in the dorsal direction. In other

9

words, each SE cell projects to topographically equivalent 3x3, 5x5 or 7x7 neighborhood of CE cells. On the ventro-dorsal orientation, the first four SE cells projected to 3x3 CE cells, the next three cells to 5x5 and finally three more cells to 7x7 cells. In order to establish planar border condition, lower frequency edge of network has been made contiguous to the high frequency end, and, due to the increasing spreading of connectivity, the network structure is not homogeneous along the isofrequency band, so that the subcortical and cortical layers, both inhibitory and excitatory, networks were doubled in a mirror image to make the transition smooth. So the simulated network has 20 by 15 cells matrices in the four subcortical and cortical layers. In the subcortical layer, each excitatory cell projects to 3x3 neighboring inhibitory (SI) cells and in turn each SI sends connection to 5x5 neighboring SE cells (see Figure 8). Similarly in the cortex, each CE cell projects to neighboring 3x3 cortical inhibitory (CI) and excitatory (CE) cells; and each CI cell is connected to 5x5 CE cells. And finally, CE cells project back to topographically equivalent 5x5 inhibitory and excitatory subcortical cells (figure 7).

We used a similar neuronal model presented previously by Grajski and Merzenich (1989). The stimulus pattern, equivalent to monotonic sound, was a triangular shaped input excitation pattern exciting 5 cells of the cochlear layer. The cell model is the same for inhibitory and excitatory cells. The cells membrane potential is given by following equation:

$$\dot{U}^i_{CE} = -Tm.U^i_{CE} + \frac{1}{M}\sum_m V^m_{SE}W^{im}_{SE\_CE} + \frac{1}{N}\sum_n V^n_{CE}W^{in}_{CE\_CE} - \frac{1}{P}\sum_p V^p_{CI}W^{ip}_{CI\_CE} \, ,$$

$$\dot{U}^i_{CI} = -Tm.U^i_{CI} + \frac{1}{M}\sum_m V^m_{CE}W^{im}_{CE\_CI} \, ,$$

$$\dot{U}^i_{SE} = -Tm.U^i_{SE} + \frac{1}{M}\sum_m V^m_{CO}W^{im}_{CO\_SE} + \frac{1}{N}\sum_n V^n_{CE}W^{in}_{CE\_SE} - \frac{1}{P}\sum_p V^p_{SI}W^{ip}_{SI\_SE} \, ,$$

$$\dot{U}^i_{SI} = -Tm.U^i_{SI} + \frac{1}{M}\sum_m V^m_{SE}W^{im}_{SE\_SI} + \frac{1}{N}\sum_n V^n_{CE}W^{in}_{CE\_SI} \, ,$$

where U^i_{XY} and V^i_{XY} are, respectively, membrane potential and cell firing rate, ranging from 0 to 100%, for post-synaptic unit/cell i of layer X with cell type Y; $Tm.$ is the membrane time constant; M, N and P are total number of input to the specific post-synaptic cell; $W^{ip}_{XY\_X'Y'}$ is the synaptic strength from the pre-synaptic p-th cell of layer X of cell type Y to the post-synaptic i-th cell of layer X' and cell type Y'. CO stands for cochlear excitatory cell. The firing rate of each cell is given by the following sigmoidal transformation of membrane potential,

$$V^i = \begin{cases} \frac{1}{2}(1+\tanh(4(U^i - \frac{1}{2}))) & U^i \geq 0.02, \\ 0 & U^i < 0.02. \end{cases}$$

The synaptic strength is updated by a) Hebbian activity-dependent changes , b) passive decay, that can be expressed by

$$W^{im}(t) = W^{im}(t-1) - \tau_{syn}W^{im}(t-1) + \alpha V^i V^m$$

$$= (1 - \tau_{syn})W^{im}(t-1) + \alpha V^i V^m \, ,$$

10

τ_{syn} is the synaptic strength decay constant and α the activity-dependent synaptic change constant for each integration step, and c) normalization, in which case the sum of all synaptic strength to cell i divided by the number of connections is kept constant to R=2 for all cells,

$$\frac{1}{N^i}\sum_j W^{ij} = R.$$

R is called the total resource of cells.

The numerical integration was carried out with a backward-Euler algorithm using integration time-steps of 0.2 unit time, τ_{syn}=0.01, Tm=0.2 and $\alpha = 0.05$.

The network was trained without supervision by applying monotonic tone inputs to the cochlea, with triangular-shaped activation of hair cells, in random frequency order. At the beginning of each stimulation cycle, the membrane potential was set to zero. then stimulus were applied for two unit time and, at the end of five unit times, synaptic strengths were updated and normalized. The training session was finished when the synaptic strength changes were smaller than 5% after 10 stimulation cycle of the complete set of 15 inputs (number of cells on frequency dimension).

Figure 9 shows the outcome of the model when stimulated with a pure tone of medium high intensity. The two left panels indicate the activity distribution of the excitatory and inhibitory cortical layers and the two panels on the right illustrate the activation of the subcortical layers. Note that the activity profile in the excitatory cortical layers shows an elevated activity along the isofrequency axis and a local minimum in the ventro-dorsal center similar to observation in the physiology. The region with a minimum excitatory activity corresponds to a region of high inhibitory activity in the subcortical layers (lower right panel). Interestingly, the cortical inhibitory layer also shows a minimum at the location of reduced excitatory cortical activity suggesting that the strength of inhibition in the center of AI may not be the only parameter that influences the degree of nonmonotonicity of cortical rate/level functions.

At this early stage of the model, several basic aspects need to be studied and improved before the outcome of the modelling can be interpreted with confidence. Among them are the following: 1) The influence of each learning and normalization parameter on the behavior of the network; 2) The use of larger networks in both dimension, so that we will be able to assess with finer detail cortical response representations; 3) Addition of more anatomical information into the model while keeping the general simplicity of the layout thus allowing reasonable margins of analyzing factors influencing the network behavior; 4) Training the network with more natural stimuli sets.

Kohonen model: Our experimental findings suggest that the global organization of auditory cortex can be described as the topographical representation of several independent stimulus domains. A factor analysis of our data indicated at least four independent parameters, namely frequency, bandwidth, intensity, and binaural interaction. Although it is possible, that other independent parameters are also represented, e.g. temporal properties, we have at this time no clear evidence regarding the nature of those parameters or their distribution pattern in AI (but see Objective 1). Based on these four identified feature dimensions, a reduction in dimensionality has to take place in order to project these four dimension onto the two spatial dimensions of auditory cortex studied so far (all of our investigations have concentrated on response properties in layers IIIb and IV). One class of models that appear to be ideally suited is the dimensionality-reducing mapping approach using Kohonen algorithms (e.g. Kohonen, 1982).

In an attempt to produce evidence that the mapping of several parameters onto the isofrequency domain of AI is compatible with principles of topology-preserving mapping emerging in Kohonen networks, we designed a simple network that projected two dimensions to a linear array. The two dimensions were selected to represent the bandwidth feature dimension and the intensity feature dimension as seen in cortical maps along the isofrequency lines. The

11

linear array corresponds to neurons along the isofrequency axis of AI. Two thousand randomized input coordinates of the two training dimensions were selected without bias for any parameter combinations. The output array had 40 elements. During the learning process, every unit of the output array (neurons in the isofrequency domain) will become maximally sensitized to a particular location in the two-dimensional input array with different units being sensitized in an orderly fashion. By plotting the output array within the original two-dimensional input plane, the general mapping can be visualized. The left column in Figure 10 shows five different outcomes of such a mapping after 32000 cycles of training (the patterns were already stable after about 5000 to 1000 cycles). Note that the shape or trajectory of the linear array across the input dimensions shows different pattern such as a circle, an S-shape, or a figure eight. Parameters such as rate of learning, number of inputs, and the number of training cycles, were kept constant. The only parameter that was changed was the size of the neighborhood that was included into the updating process. The largest neighborhood size was 39 (upper panel) and the smallest neighborhood was 2 (lowest panel). Note the increase in complexity of the trajectories for the smaller neighborhood sizes.

In order to allow a more direct comparisons with the experimental data, the projections of the trajectory on the two parameter axes is plotted in the right column of Figure 10. The dotted lines show the change of the bandwidth factor along the hypothetical isofrequency domain and the solid line shows the spatial variation of the intensity factor. Two aspects of these single factor trajectories should be noted. First, the topology-preserving mapping results in fairly smooth and nearly sinusoidal variations of each parameter through its parameter space. However, the spatial frequency of this parameter sequence can change. This is indicated by the different spatial frequency estimates noted for each trajectory. The spatial frequency values for the intensity factor (SF(I)) vary between 1 cycle along the isofrequency domain (upper diagram) to 2 cycles (lower diagram). In this case the spatial frequency of the bandwidth factor (SF(B)) varied between 0.5 and 1 cycles. Second, despite the smooth and approximately sinusoidal course of the parameter trajectories, essential all trajectories show some regions of little change of one or both of the two parameters, resulting in regions with fairly similar parameter values. Both of these effects, fairly smooth general gradients of parameters along the isofrequency lines with spatial frequencies of the same order of magnitude and small regions with a constant parameter value, are also characteristic for the parameter distribution actually observed in AI. Figure 11 shows the actual trajectories of bandwidth and intensity factors along the isofrequency axis of five cats. Although these real data are more noisy than the model maps, similarities to the model results can be noted. Some two-dimensional trajectories show an elliptical shape, other resemble a figure eight. Accordingly, the trajectories plotted for each parameter separately show different spatial frequencies. The actual spatial frequencies are between 0.5 and 2.5 not unlike the range emerging from the model. In addition, the real parameter distributions show a 'patchiness', or regions of little change in a given parameter. This patchiness has been used by others to argue against the existence of real parameter gradients in the organization of the isofrequency domain. From the modeling results it appears that general gradients and local patches are both necessary consequences of an selforganizing, dimension-reducing, topology-preserving algorithm.

The input distributions of the parameters to the Kohonen network covered all parameter combinations with equal probability. This is not necessarily the case for the real auditory system. In fact, the course of some of the trajectories (oblique ellipses) indicate that certain parameter combinations appear to be slightly overrepresented. Whether this is a consequence of the actual input probability distributions or might be related to the existence of other parameters that are also projected onto this space and may distort the representation of these two dimension remains to be seen.

It is concluded that selforganizing Kohonen networks are in reasonable agreement with actually observed cortical parameter representations and that smooth parameter gradients with occasional patchiness is a natural consequence of topology preserving algorithms.

Summary:

Combined, the cortical experiments and modelling attempts represent an approach to bioacoustic signal classification that is based on the biological transformation and parametric representation of complex signals in the mammalian CNS of animals with 'open reception', similar to humans. Application of similar principles to technical signal classification schemes should result in an improvement of the classification success since the utilized mechanisms contain biological proven aspects that aid the animal in a robust signal classification. The parallel electrophysiological and modelling study of the cortical representation of signals in noisy and quiet background have revealed mechanisms that are responsible for the pre-processing in successful classification schemes and can enhance our understanding of the signal analysis properties and their implementation in the CNS.

References:

G.D. Angelis, I. Ohzav'a, R.D. Freeman, J.Neurophys. 69:1091-1117, 1993.

H. R. Dinse and C.E. Schreiner, Soc. Neurosci. Abstr. 18:150, 1992.

H.R. Dinse, C.E. Schreiner, F. Spengler, B. Godde, B. Hartfiel (in press, see Appendix)

J.J. Eggermont, Hear. Res. 56:153-167, 1991.

J.J. Eggermont , P.I.M. Johannesma, A.M.H.J. Aersten, Q. Rev. Biophys. 16:341-414, 1983.

G. Ehret, M.M. Merzenich, Brain Res. Revs. 13:139-163, 1988.

K.A. Grajski and M.M. Merzenich, In: Neural information processing systems, Vol 2., ed. D.E.
    Tourefsky, Morgan Kaufman, San Mateo (1990).

D.A. Hillier, Doctoral Thesis, Washington University:St. Louis, 1991.

J.P Jones, L.A.Palmer, J. Neurophysiol. 58:1187-1211, 1987.

T. Kohonen, Biol. Cyb. 43:59-69, 1982.

J. Mendelson, C.E. Schreiner, M. Sutter, K. Grasse, Exp. Brain Res.94:65-87, 1993.

M.M. Merzenich, C. Schreiner, W. Jenkins, X. Wang, Annals of the New York Academy of
Sciences 682: 1-22, 1993.

M. Porat, Y.Y. Zeevi, IEEE Transact. Biomed. Eng. 36:115-129, 1989.

D.P. Phillips, Behav. Brain Res. 37:197-214, 1990.

G. Recanzone, C.E. Schreiner, M.M. Merzenich, J. Neurosci. 13:87-103,1993.

C.E. Schreiner, J. Mendelson, M. Sutter, Exp. Brain Res. 92:105-122, 1992.

C.E. Schreiner and M. Sutter, J. Neurophys. 68:1487-1502, 1992.

C.E. Schreiner and J.V. Urbas, Hear. Res.21:227-241, 1986.

C.E. Schreiner and J.V. Urbas, Hear. Res.32:49-64, 1988.

S. Sharnma., J. Fleshman,P. Wiser, H. Versnel, J. Neurophysiol.69:367-383, 1993.

P. Tallal, S. Miller R.H. Fitch, Annals of the New York Academy of Sciences 682:27-47, 1993.

Figure Legends:

Figure 1: Reverse correlation of neuronal responses in auditory cortex with input signal. A: Pre-event stimulus ensemble (PESE) of a cortical neuron in response to vocalizations. A time interval of 30ms before the generation of an action potential is shown. The signals that preceed each spike (N>100) have been averaged in 128 frequency channels after lowpass filtering each channel. The resulting temproal envelop average is proportional to the probability that a certain stimulus configuration has to occur before a spike can be elicited. Red: probability high above average; green: average probability, blue: probability below average. B:PESE for pure tone stimulation. Same convention as for panel A. C: Color inverted PESE for vocalizations (panel A). blue: probability high above average; green: average probability, red: probability below average. D: Frequency response area of neuron. 675 different tones were used to generate the intensity-frequency response profile. Red: evoked response.

Figure 2: Reverse correlation from neuronal responses in auditory cortex.Same convention as in Figure 1.

Figure 3: Modelled neuronal activity profile across auditory cortex. The neuronal activity (realtive spike rate) is plotted as elevation. The other two axis correspond to the isofrequency axis, represented by 133 neurons and the frequency axis represented in octaves. Low numbers along the isofrequency axis correspond to the dorsal portion of AI. The stimulus was a pure tone with a frequency corresponding to the center of the frequency axis and an intensity of 10dB.

Figure 4: Modelled neuronal activity profile across auditory cortex. The stimulus was a pure tone with a frequency corresponding to the center of the frequency axis and an intensity of 45dB. Same conventions as in Figure 3.

Figure 5: Modelled neuronal activity profile across auditory cortex. The stimulus was a pure tone with a frequency corresponding to the center of the frequency axis and an intensity of 45dB. Same conventions as in Figure 3.

Figure 6: Connectivity organization in cortical model.The first layer (cochlear input) consists of a one dimensional array of 15x1 cells representing the cochlear band-pass filter and hair cells. The subcortex has two layers, one excitatory and one inhibitory, of 15x10 cells. The cortex is represented by the same number of 15x10 excitatory and inhibitory cells. Each cell of the cochlear layer projects to a band of 3 by 10 subcortical excitatory (SE) cells preserving the isofrequency representation along one dimension of subcortical layer.The spread of connections (light shade) away from the isofrequency axis is uniform in the subcortical layer and not uniform in the cortical layer.

Figure 7: Each SE cell projects to topographically equivalent 3x3, 5x5 or 7x7 neighborhood of CE cells. On the ventro-dorsal orientation, the first four SE cells projected to 3x3 CE cells, the next three cells to 5x5 and finally three more cells to 7x7 cells.

Figure 8: In the subcortical layer, each excitatory cell projects to 3x3 neighboring inhibitory (SI) cells and in turn each SI sends connection to 5x5 neighboring SE cells.
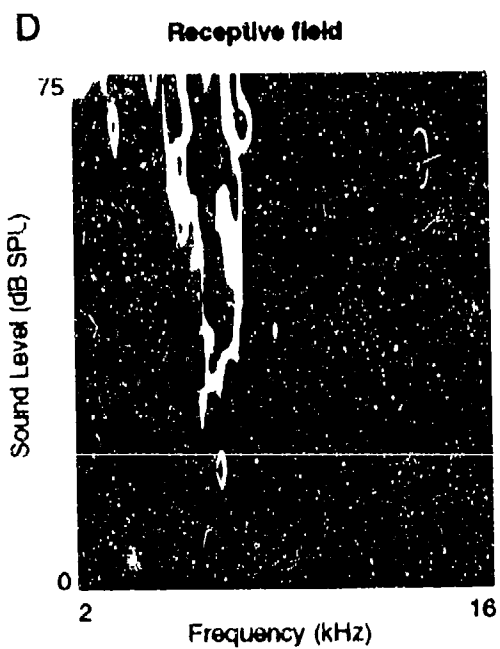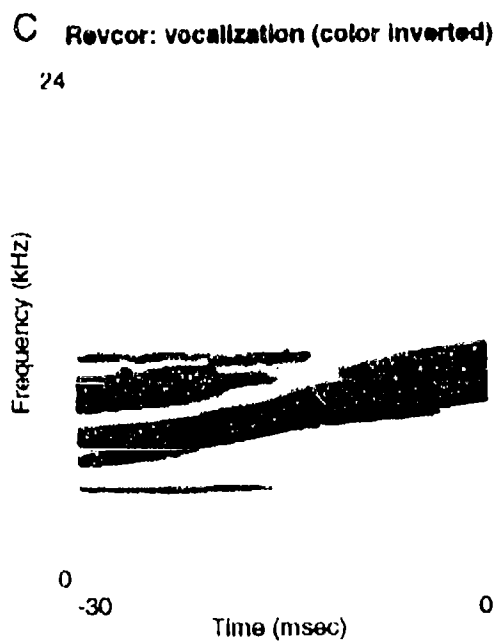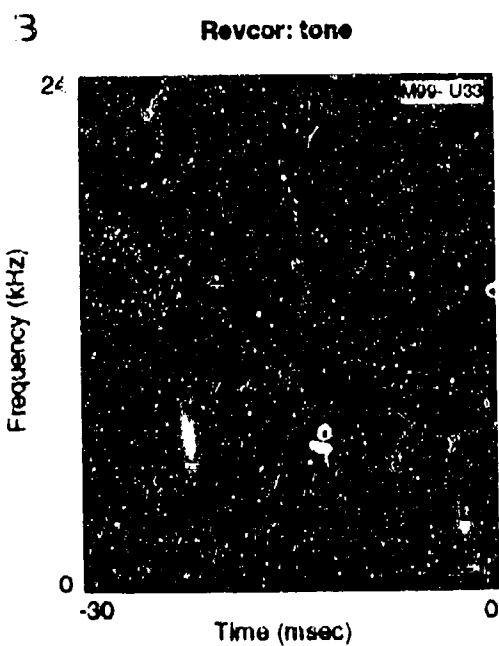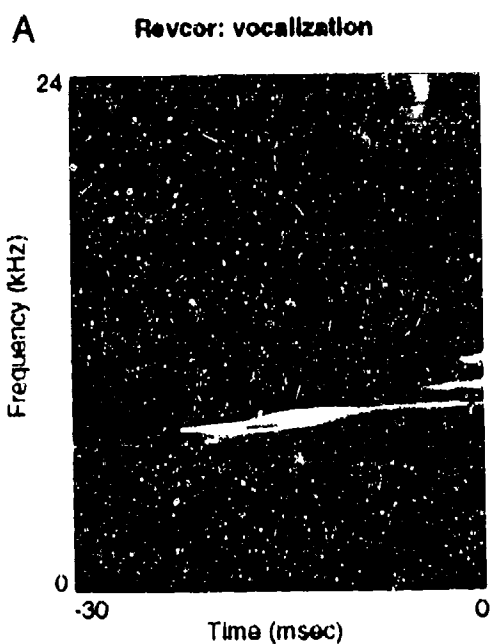
Figure 9: Activity distribution in subcortical and cortical layers when stimulated with tone of moderate intensity. The two left panels indicate the activity distribution of the excitatory and inhibitory cortical layers and the two panels on the right illustrate the activation of the subcortical layers.

Figure 10: Cortical dimension reduction and spatial frequency emergence based on Kohonen mapping. A two-dimensional array was projected onto a linear array using a Kohonen algorithm.

15

The left column shows the results of five different runs of the model. Each of the linear array neurons (N=40) is plotted at the location of the original input space that provides the strongest input (strongest weight). Parameter is the size of the neighborhood included in the updating of the weights. Upper panel: Neighborhood size = 39, lowest panel: Neighborhood size:2. In the right columns, the trajectories are plotted for each input dimension separately. Dashed line: bandwidth dimension; solid line: intensity dimension. The spatial frequency (SF) was obtained for each curve by estimating the lowest Fourier component that would describe each trajectory.

Figure 11: Actual trajectories in bandwidth-intensity space for the isofrequency domain (10-12kHz) of five animals. The actual trajectories was extracted form the data by reducing the local parameter scatter through a moving average window of length 10. Conventions same as in Figure 10.
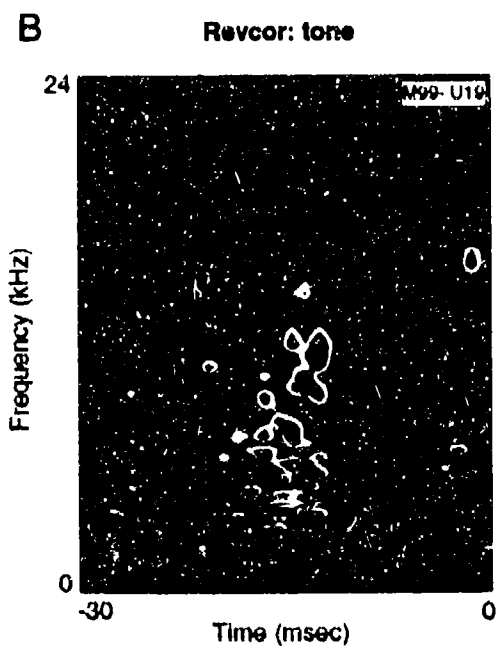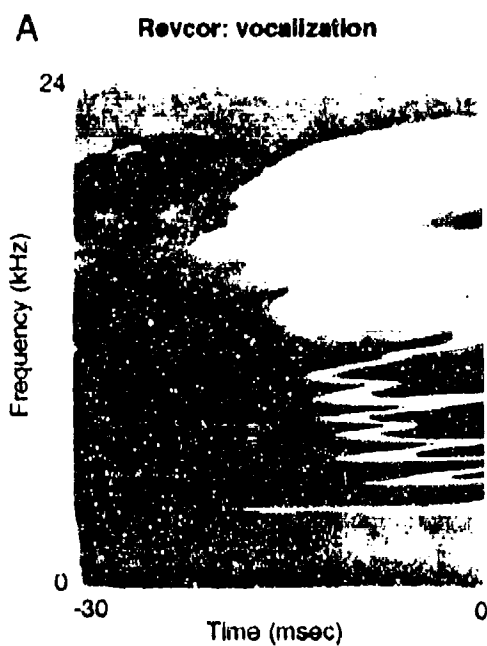
Appendix:

A Revcor: vocalization

B Revcor: tone

M00- U33

C Revcor: vocalization (color inverted)

D Receptive field

A   Revcor: vocalization

B   Revcor: tone

C   Revcor: vocalization (color inverted)

D   Receptive field

Figure 3

10dB

relative spike activity

100
50

20
40
60
80
100
120
-2

neuron

2
1
0
-1

frequency / octaves

Figure 4

45dB

Figure 5

90dB

Figure 6

Caudal

Dorsal

Rostral

Ventral

Cortical
Layer
(AI)

Subcortical
Layer

Low
Freq

High
Freq

Cochlear
Input
(from Pre-Processor)

Figure 7

Cortical Layer (AI)

Subcortical Layer

Cochlear Input
(From Pre-Processor)

Figure 8

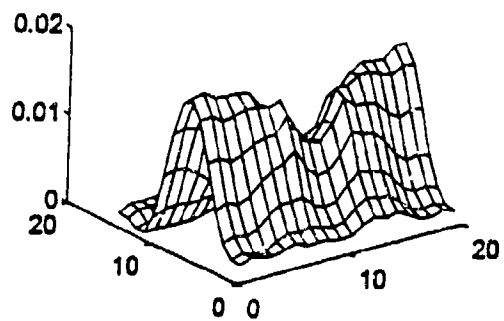Cortical Layer (AI)

Subcortical Layer

Figure 9

Uce

Use

Uci

Usi

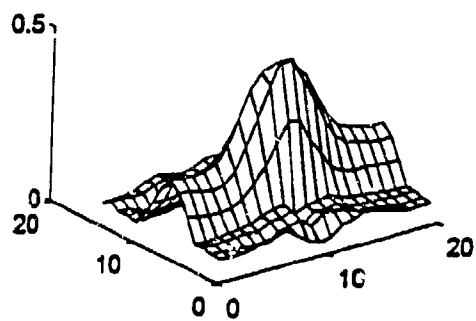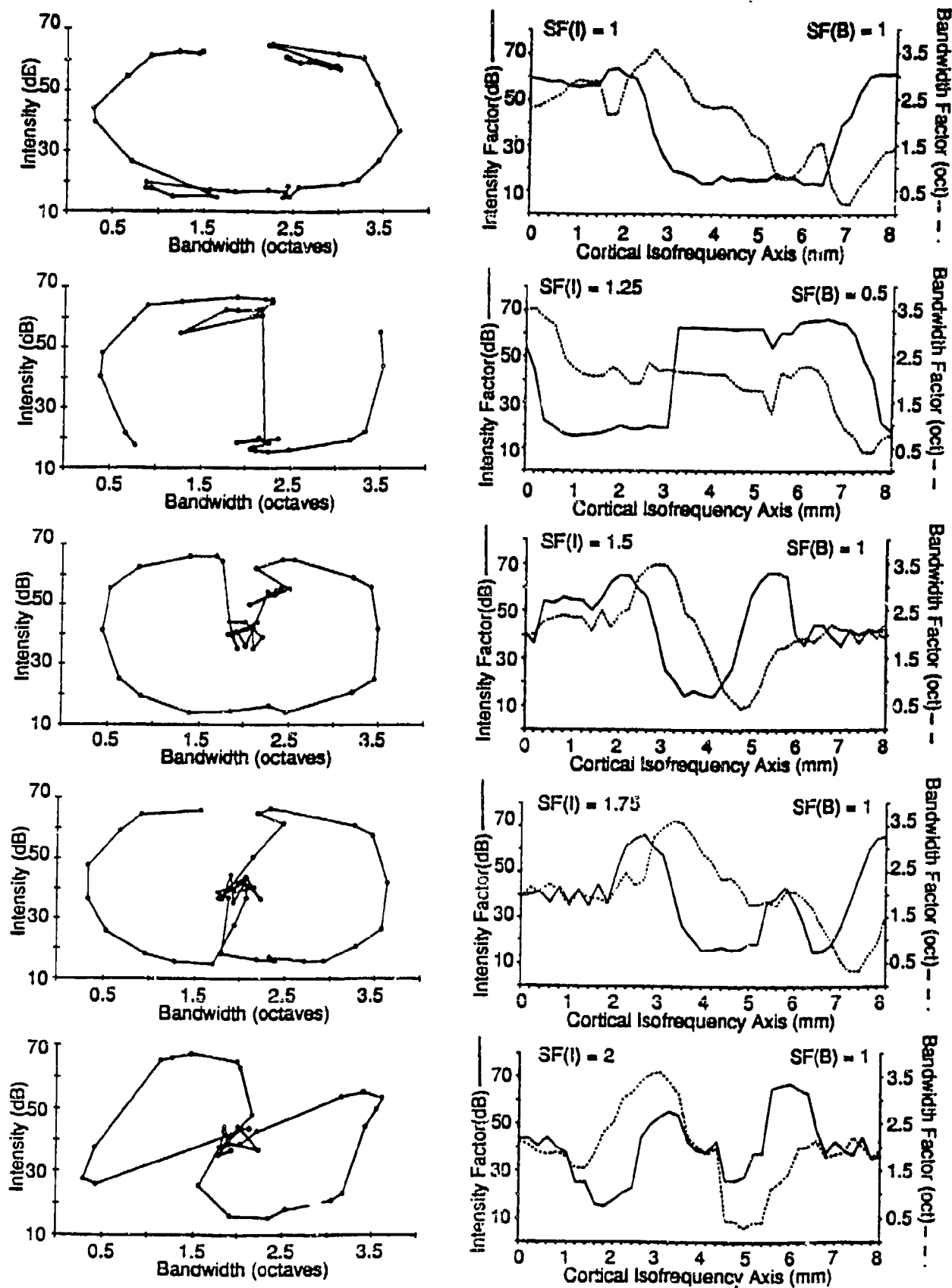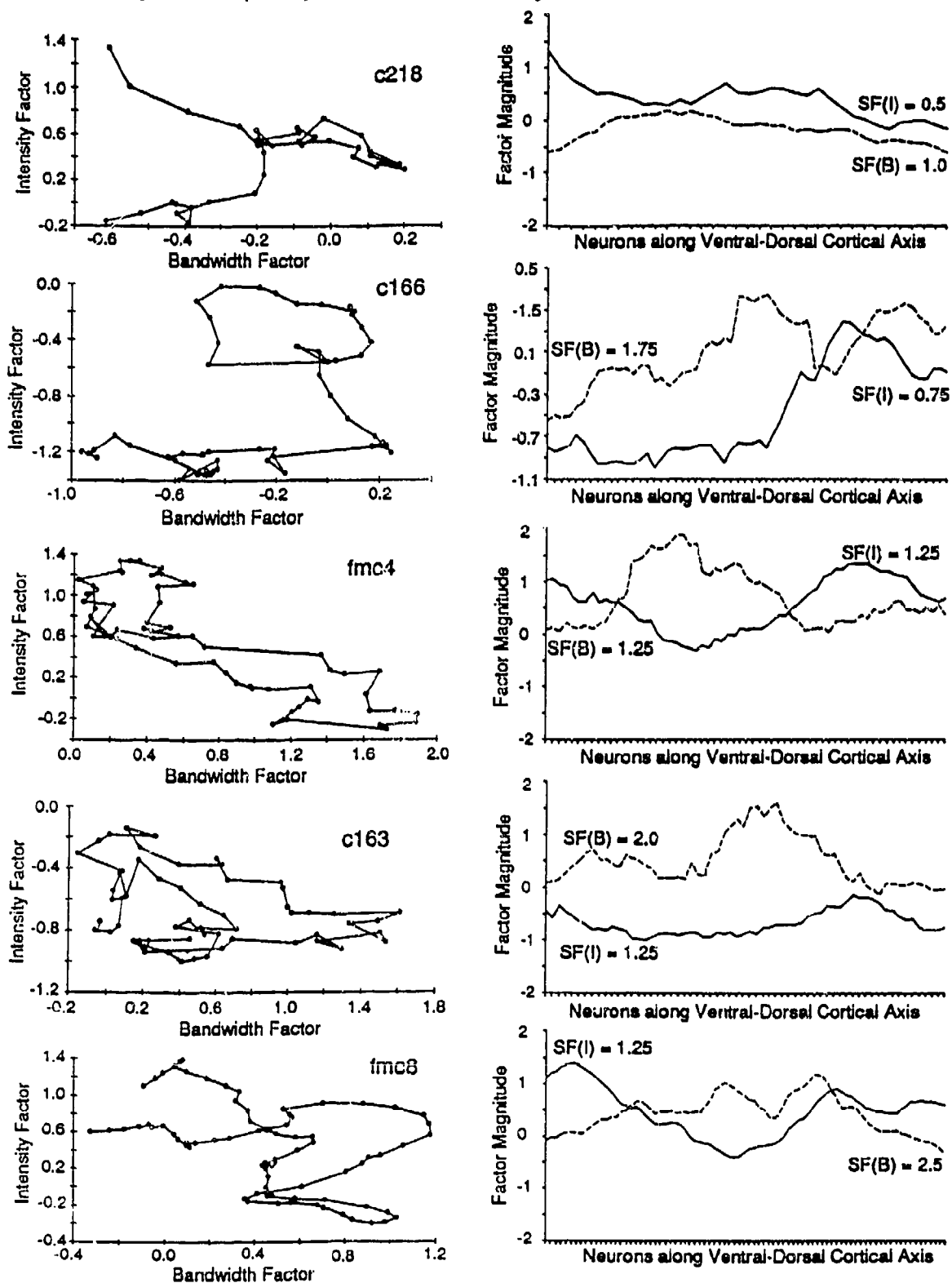Figure 18

Cortical Dimension Reduction and Spatial Frequency Emergence
based on Kohonen Mapping

Figure 11
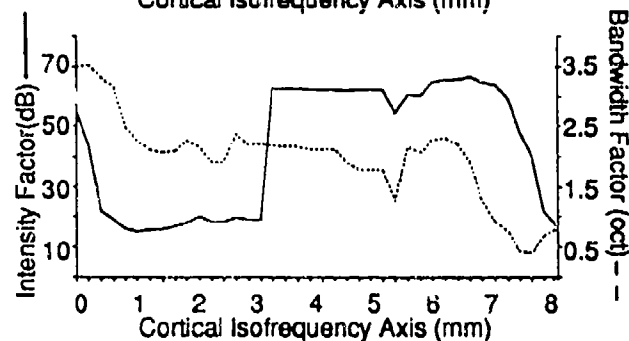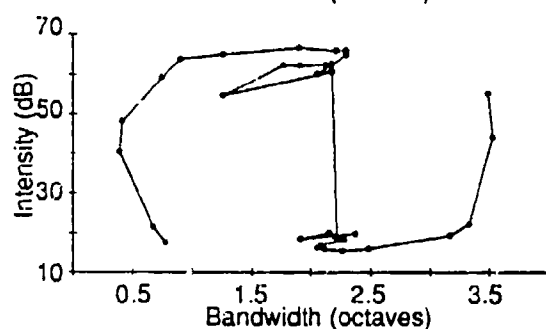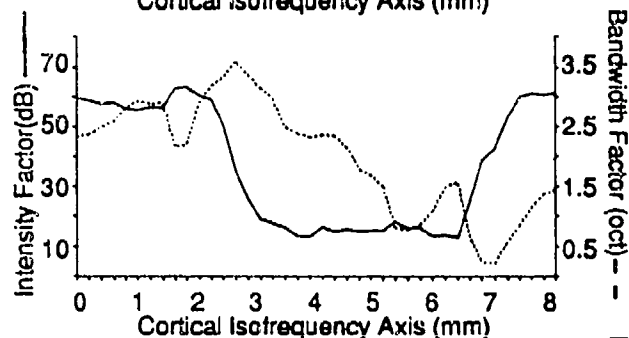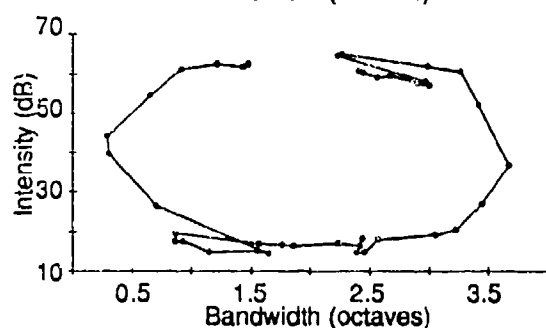
Spatial Frequency Estimate for Intensity and Bandwidth Factors in AI
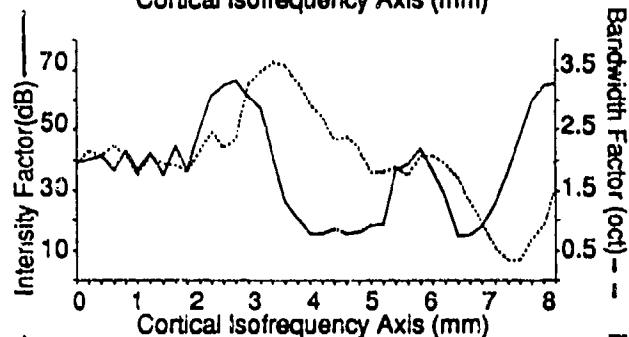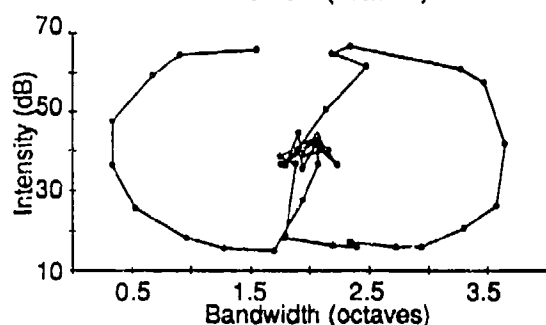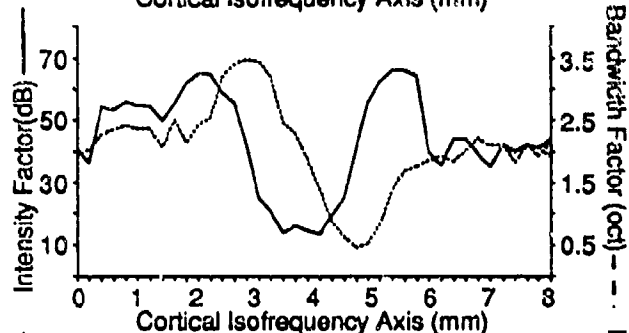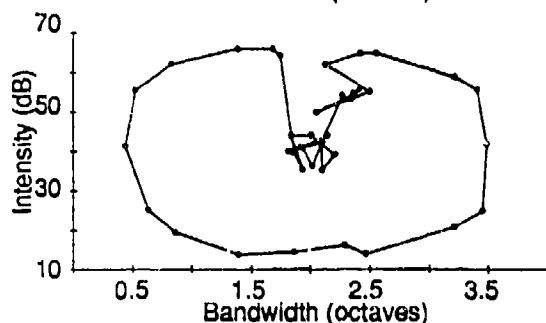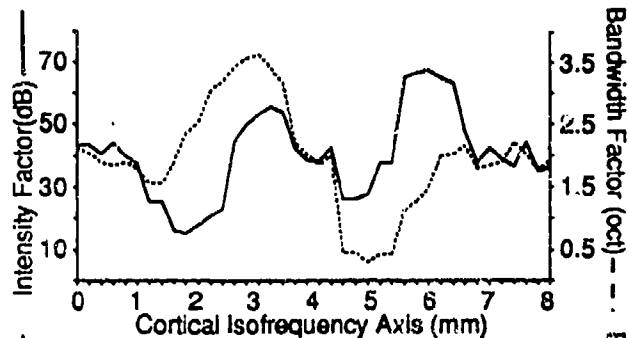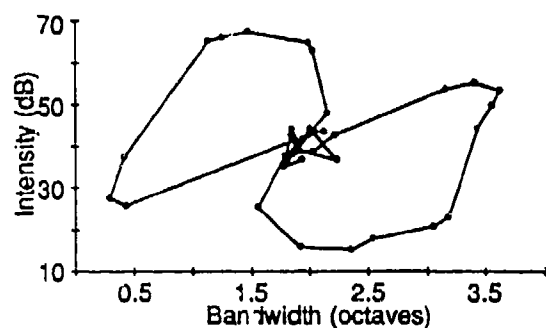
Topography of Excitatory Bandwidth in Cat Primary Auditory Cortex: Single-Neuron Versus Multiple-Neuron Recordings

CHRISTOPH E. SCHREINER AND MITCHELL L. SUTTER
*Coleman Memorial Laboratory, W. M. Keck Center for Integrative Neuroscience, University of California
at San Francisco, San Francisco, California 94143-0732*

SUMMARY AND CONCLUSIONS

1. The spatial distribution of the sharpness of tuning of single neurons along the dorsoventral extent of primary auditory cortex (AI) was studied. A sharpness of tuning gradient was initially obtained with multiple-unit recordings, and in combination with the cochleotopic organization, served as a frame of reference for the locations of single neurons. The frequency selectivity or "integrated excitatory bandwidth" of multiple units varied systematically along the dorsoventral extent of AI. The most sharply tuned unit clusters were found at the approximate center of the dorsoventral extent. A gradual broadening of the integrated excitatory bandwidth in both dorsal and ventral directions was consistently seen.

2. The multiple-unit measures of the bandwidth 10 (BW10) and 40 dB (BW40) above minimum threshold, pooled across several animals and expressed in octaves, were similar to those described within individual cases in cats. As in the individual animals, the bandwidth maps were V shaped with minima located at the approximate center of the dorsal-ventral extent of AI. The location of the minimum in the multiple-unit bandwidth map (i.e.. the most sharply tuned area) was used as a reference point to pool single-neuron data across animals.

3. For single neurons, the dorsal half of the BW40 distribution showed a gradient paralleling that found for multiple units. For both single and multiple units, the average excitatory bandwidth increased at a rate of ~0.27 octaves/mm from the center of AI toward the dorsal fringe. Differing from the dorsal half of AI, the ventral half of AI showed no clear BW40 gradient for single units along its dorsoventral extent. At 40 dB above minimum threshold, most ventral neurons encountered were sharply tuned. By contrast, the multiple-unit BW40 showed a gradient similar to the dorsal half with 0.23 octaves/mm increasing from the center toward the ventral border of AI.

4. For single neurons, BW10 showed no clear systematic spatial distribution in AI. Neither the dorsal nor the ventral gradient was significantly different from zero slope, although the dorsal half showed a trend toward increasing BW10s. Contrasting single neurons, both dorsal and ventral halves of AI showed BW10 slopes for multiple units confirming a V-shaped map of the integrated excitatory bandwidth within the dorsoventral extent of AI.

5. On the basis of the distribution of the integrated (multiple-unit) excitatory bandwidth, AI was parceled into three regions: the dorsal gradient, the ventral gradient, and the central, narrowly tuned area. In ventral AI, single units were significantly more sharply tuned than multiple units for BW10 and BW40. In dorsal AI, single units were not statistically different from multiple units for BW40. In central AI, single units were significantly sharper for BW40, but not BW10.

6. Estimates of the scatter of characteristic frequency (CF) of single neurons in the dorsoventral extent domain were obtained relative to the frequency organization determined with multiple units. The central narrowly tuned region showed the least CF scat-

ter. The CF scatter increased toward the dorsal, and in particular, the ventral end of the dorsoventral extent of AI.

7. The combined single- and multiple-unit results suggest that AI is composed of at least two functionally distinct subregions along the dorsoventral extent/isofrequency domain on the basis of the bandwidth properties of tuning curves. The dorsal region (AId) displays a global gradient of BW40 expressed in single- and multiple-unit measurements, and contains broadly and sharply tuned single-peaked as well as multipeaked neurons. The ventral region (AIv) predominantly contains neurons narrowly tuned at 40 dB above threshold. The increase in integrated (multiple-unit) excitatory bandwidth in AIv can be related to a progressively larger CF scatter at more ventral locations. The transition between AId and AIv is delineated by a reversal in the BW40 gradient for multiple units coinciding with a region of small CF scatter and sharp tuning of the underlying single neurons.

INTRODUCTION

Recent studies of cat auditory cortex have demonstrated topographic order within the electrophysiologically determined organization of the dorsoventral extent of the primary auditory cortical field (AI) (Mendelson et al. 1988; Schreiner and Cynader 1984; Schreiner and Mendelson 1990; Schreiner et al. 1988; Sutter and Schreiner 1991a) in addition to previously described organizations relating to binaural response properties (Imig and Adrian 1977; Imig et al. 1990; Middlebrooks et al. 1980; Rajan et al. 1990; Reale and Kettner 1986). Several response parameters, including sharpness of frequency and amplitude tuning, obtained from multiple-unit recordings, were found to be nonrandomly distributed along the dorsoventral dimension of AI. These findings suggest a systematic functional organization in the isofrequency domain of AI that involves the coding of basic signal properties, among them signal intensity and spectral complexity (Schreiner and Mendelson 1990; Shamma and Fleshman 1990; Sutter and Schreiner 1991a). The functional interpretation of topographies based on multiple-unit recordings is, however, limited and can only represent an approximation of actual cortical processing. Ultimately, the properties of single neurons underlying those integrated measures of cortical activity must be analyzed and related to multiple-unit findings to better understand the basic principles and spatial distribution of physiological characteristics and functional organization of AI.

Studying the spatial distribution of single-neuron response properties across large portions of AI has proven to be a difficult task for a number of reasons. Historically,

single-unit studies that used anatomically based pooling strategies (e.g., Evans and Whitfield 1964; Goldstein et al. 1970) failed to fully support basic organizational principles such as a strong cochleotopicity of AI as obtained with integrated responses. For example, using sulcal patterns and vasculature as landmarks for pooling topographical data across animals is untenable because both display a large variability among animals (Merzenich et al. 1975). Additionally, the cytoarchitectonic boundaries of AI are not of sufficient sharpness (e.g., Rose 1949; Winer 1984) to allow a precise alignment of locations within a given field from different animals. Characterizing a sufficient number of single neurons in one experiment to construct a reliable physiological map is difficult in a cortical field of the extent of cat AI.

An alternative to utilizing anatomic landmarks is the use of known systematic spatial distributions of physiological features that may provide information for the proper spatial alignment of fields from several individual cortices. The most basic choice is the use of the cochleotopic frequency gradient, reliably determinable with multiple-unit recordings (e.g., Merzenich et al. 1975; Reale and Imig 1980; Schreiner and Mendelson 1990), to determine the rostrocaudal extent of the field and to identify locations within this dimension. However, a second parameter is necessary to align locations along the dorsoventral extent of AI, i.e., approximately along its isofrequency domain. The clustered distribution of binaural response characteristics along the dorsoventral extent of AI is quite variable from animal to animal (Imig and Adrian 1977; Imig and Brugge 1978; Imig and Reale 1981; Middlebrooks et al. 1980; Reale and Kettner 1986; Schreiner and Cynader 1984) and, therefore, seems not to be suitable as a basis for a spatial normalization or alignment of AI from different cortices.

A second physiological parameter that may be particularly useful in aligning cortical fields is the spatial distribution of the integrative excitatory bandwidth along the dorsoventral dimension of AI. A recent study (Schreiner and Mendelson 1990) revealed a region of sharp frequency tuning of multiple-unit responses within the central portion of the isofrequency domain contiguous with ventral and dorsal gradients of decreasing frequency selectivity. This spatial distribution of the integrated excitatory bandwidth in AI (Schreiner and Cynader 1984; Schreiner and Mendelson 1990) appears to be *1*) similar and less variable from animal to animal and *2*) appears to be consistent over a wide range of frequencies. Therefore it may serve as a topographical frame of reference for comparing and pooling data from different animals. In a previous report this approach was successfully utilized to determine the spatial distribution of a subpopulation of auditory cortical neurons, namely neurons with multipeaked tuning curves, in the dorsoventral dimension of AI (Sutter and Schreiner 1991a). By pooling data across animals, a sufficient number of topographically identifiable neurons with multipeaked tuning curves was obtained to derive a statistically secured estimation that their locations were essentially confined to the dorsal part of AI. This paper explores the property and topography of frequency selectivity of single-neuron responses along the dorsoventral extent, approximating the isofrequency axis of cat AI. Single-neuron

positions were obtained relative to gradients in integrated excitatory bandwidth determined with multiple-unit recordings. Some results have previously been presented in abstract form (Sutter and Schreiner 1991b).

METHOD

Surgical preparation

The methods are similar to those described in Sutter and Schreiner (1991a). Experiments were conducted on nine young adult cats. Anesthesia was induced with an intramuscular injection of ketamine hydrochloride (10 mg/kg) and acetylpromazine maleate (0.28 mg/kg). After venous cannulation an initial dose of pentobarbital sodium (to effect, ~30 mg/kg) was administered. Animals were maintained at a surgical level of anesthesia with a continuous infusion of pentobarbital sodium (2 mg·kg·h) in lactated Ringer solution (infusion volume, 3.5 ml/h) and, if necessary, with supplementary intravenous injections of pentobarbital sodium. The cats were also given dexamethasone sodium phosphate (0.14 mg/kg im) to prevent brain edema, and atropine sulfate (1 mg im) to reduce salivation. The temperature of the animals was recorded with a rectal temperature probe and maintained at 37.5°C by means of a heated water blanket with feedback control.

The head was fixed, leaving the external meati unobstructed. The temporal muscle over the right hemisphere was then retracted and the lateral cortex exposed by a craniotomy. The dura overlaying the middle ectosylvian gyrus was removed, the cortex was covered with silicone oil, and a photograph of the surface vasculature was taken to record the electrode penetration sites. For recording topographically identified single neurons, a wire mesh was placed over the craniotomy, and the space between the grid and cortex was filled with a 1% solution of clear agarose. This approach diminished pulsations of the cortex and provided a fairly unobstructed view of identifiable locations across the exposed cortical surface.

Stimulus generation and delivery

Experiments were conducted in a double-walled sound-shielded room (IAC). Auditory stimuli were presented via a sound delivery system designed to provide essentially flat transfer function when connected to the average cat ear (Sokolich 1981, US Patent 4251686). Headphones (STAX 54) enclosed in small chambers were connected to sound delivery tubes sealed into the acoustic meati. The sound delivery system was calibrated in vitro with a sound level meter (Brüel & Kjaer 2209) and a waveform analyzer (General Radio 1521-B). The frequency response of the system was essentially flat up to 14 kHz and did not have major resonances deviating more than ±6 dB from the average level. Above 14 kHz, the output rolled off at a rate of 10 dB/octave. However, the actual in vivo transfer function may have deviated from this estimate. Harmonic distortion was below 0.18% (−55 dB) for the maximum input voltage to the speakers of 0.35 V_{RMS}. No influences of harmonic distortion on the obtained tuning curves was ever noted.

Tones were generated by a microprocessor (TMS32010: 16-bit D/A converter at 120 kHz sampling rate; low-pass filter of 96 dB/octave at 15, 35, or 50 kHz). The amplitude of the signals was controlled by varying the number of amplitudes steps in generating the waveform (each step corresponded to ~0.15 mV). The highest number of amplitude steps was 63,246 and the lowest number was 20, resulting in a useful dynamic range of 70 dB. Additional attenuation was provided by a pair of passive attenuators (Hewlett Packard). The duration of the tone bursts was usually 50 ms (tone bursts were extended to 85 ms for long-latency re-

sponses) including 3-ms rise/fall time. The interstimulus interval was 400–1.000 ms.

Frequency response areas (FRAs)

FRAs were obtained for each neuron. To generate an FRA, at least 675 different tone bursts were delivered. Tone bursts were presented in a pseudorandom sequence of different frequency/level combinations selected from 15 level values and 45 frequency values. Steps between levels were 5 dB, resulting in a sampled dynamic range of 70 dB.

The frequency range covered by the 45 frequency steps was centered around the estimated characteristic frequency (CF) of the recording site and covered between 2 and 5 octaves, depending on the width of the frequency tuning curve as obtained by audiovisual criteria. Stimulus frequencies were chosen so that the 45 presented frequencies were spaced an equal fraction of an octave over the entire range (for most cases this provided a 0.067-octave resolution over a total of 3 octaves).

Recording procedure

Parylene-coated tungsten microelectrodes (Microprobe) with impedances of 1.0–8.5 MΩ at 1 kHz were introduced into the auditory cortex with a hydraulic microdrive (Kopf) remotely controlled by a stepping motor. All penetrations were essentially orthogonal to the brain surface. The recordings reported here were derived at intracortical depths ranging from 600 to 1.000 μm as determined by the microdrive setting, corresponding to portions of cortical layers 3 and 4. Neuronal activity of single neurons, and for the initial mapping, small groups of neurons were amplified, band-pass filtered, and monitored on an oscilloscope and an audio monitor. Action potentials were isolated from the background noise with a window discriminator (BAK DIS-1). The number of spikes per presentation and the arrival time of the first spike after the onset of the stimulus were recorded and stored in a computer (DEC 11/73). The recording window had a duration of 50–85 ms, corresponding to the stimulus duration and excluding any offset responses.

Data analysis

From the responses to 675 different frequency/level combinations, an objectively determined FRA was constructed for every neuron or group of neurons. Figure 1 shows an example of three reconstructed FRAs obtained in AI. The ordinate corresponds to the sound level of the tone-burst stimulus, whereas the abscissa corresponds to the frequency. All presented stimuli would be represented by a 15 (ordinate) by 45 (abscissa) grid with equal spacing and size of elements. Responses are represented at the point of intersection of the intensity and frequency of each presented stimulus by a small filled rectangle. The vertical length of each rectangle is proportional to the number of spikes discharged in response to the stimulus. The absence of a rectangle for a given point of stimulus presentation represents no response. Usually each stimulus was presented once. If the resulting FRA was not well defined, the process was repeated with the same 675 stimuli, and the resulting evoked activity was added to the first set of responses.

A tuning curve was extracted from the FRA by the use of an objective method. For the response threshold, a computer program defined the isoresponse criteria as the spontaneous rate (estimated from 45 points outside the response area) plus 20% of the peak rate (9-point weighted averaging). Weighted 9-point smoothing/averaging was used to determine the response at each point because it provided a more reliable estimate by increasing the effective number of repetitions per location at the expense of some frequency and intensity resolution (Sutter and Schreiner
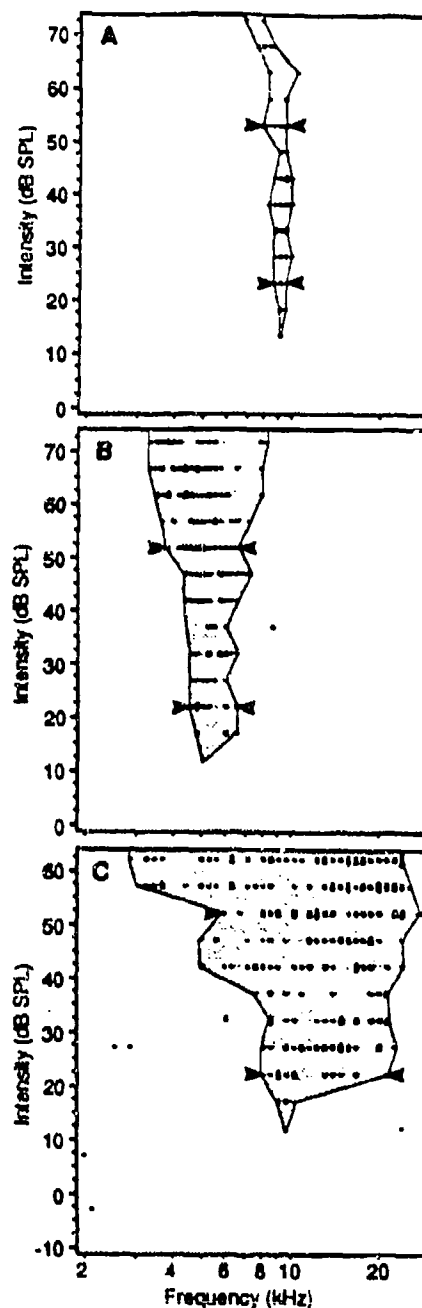
FIG. 1. Frequency response areas (FRAs) for 3 neurons in AI. A total of 675 tone bursts with different frequency/level combinations were presented. Frequency/level combinations marked with filled rectangles elicited action potentials. Height of each rectangle corresponds to number of action potentials. FRAs are plotted for a frequency range of 4 octaves. Bandwidth values were obtained 10 (open arrows) and 40 dB (filled arrows) above minimum threshold. A: characteristic frequency (CF) = 9.05 kHz, BW10 = 0.13 octaves (Q-10dB = 11.3), BW40 = 0.26 octaves (Q-40dB = 5.66). B: CF = 5 kHz, BW10 = 0.55 octaves (2.38), BW40 = 0.80 octaves (1.79). C: CF = 9.7 kHz, BW10 = 1.44 octaves (0.70), BW40 = 2.12 octaves (0.49).

1991a). This criterion was robust, yielding comparable tuning curves for the wide range of FRAs recorded, including high spontaneous multiple-unit and low spontaneous single-unit recordings. Shaded areas in Fig. 1 correspond to the objectively determined FRAs; the border of the shading corresponds to the tuning curve. Several response properties were measured from each FRA.

Among them were *1*) the characteristic frequency (CF; the stimulus frequency with the lowest sound pressure level necessary to evoke neuronal activity); *2*) BW10 (the bandwidth of the FRA, in octaves, 10 dB above minimum threshold), and *3*) BW40 (the bandwidth of the FRA, in octaves, 40 dB above minimum threshold).

For units with multipeaked tuning curves, the entire response area including the nonresponsive area between individual peaks was used to determine the excitatory bandwidth.

RESULTS

Results are based on a total of 146 multiple units and 103 single neurons from 9 cats. On the average, multiple-unit maps were derived from ~15 fairly evenly spaced recording locations over a restricted CF range (usually <2 octaves). From the multiple-unit recordings, three parameters were extracted. First, the AI/AII border was determined in accordance with criteria suggested by Schreiner and Cynader (1984). The criteria for AII are based on a reliable and spatially consistent shift of three multiple-unit physiological properties: *1*) an increase in thresholds by ≥15 dB, *2*) Q-10dB values (CF/excitatory bandwidth 10 dB above minimum threshold) below a certain CF-dependent value, and *3*) a blurring of the CF topography. Single neurons falling into AII were not included in the analyzed set of data. Second, the approximate orientation of the isofrequency contours within AI was determined, and third, the distribution of the integrated excitatory bandwidth 10 and 40 dB above minimum threshold was reconstructed.

A large range of bandwidths for multiple and single units were encountered in this series of experiments corresponding to ranges seen for Q-10dB and Q-40dB for multiple units in previous experiments (Schreiner and Mendelson 1990), as well as for Q-10dB (Phillips and Irvine 1981) and bandwidth values (Abeles and Goldstein 1970a,b; Hind 1960) for single neurons. Figure 1 displays typical examples of FRAs with sharp, medium, and broad frequency tuning of single-neuron responses. Bandwidth measures BW10 and BW40 for the units are indicated.

Topography of integrated excitatory bandwidth

To characterize topographically the dorsoventral axis of AI, electrode penetrations orthogonal to the cortical surface were placed, in most cases, in close approximation to the orientation of isofrequency contours. The actual CFs used to create a map are illustrated for two examples in Fig. 5. For most cases, isofrequency contours in AI could be approximated with a straight line deviating slightly from the dorsal-ventral direction (see Fig. 5C, the isofrequency axis is roughly parallel to the abscissa). Occasionally, the angle between the orientation of the approximated course of an isofrequency contour deviated as much as 30° from the dorsoventral axis. The CF topography became weaker at the most dorsal extent of the ectosylvian gyrus where isofrequency contours often slanted posteriorly (see Fig. 5, *A* and *B*) (Middlebrooks and Zook 1983; Sutter and Schreiner 1991a). Cortical recording locations are represented as dorsoventral distance along the approximated orientation of the straight portion of isofrequency contours. The range of CFs included for each case varied between 0.2 and 3 octaves. Essentially all locations included in the current data

evaluation had CFs >4 kHz. In contrast to most previous studies, the sharpness of frequency tuning in this study was measured as bandwidth expressed in octaves. Figure 2 displays BW10 and BW40 for all single neurons in this study. Similar to the traditionally used Q-factor (CF/bandwidth), the bandwidth measure reduces the influence of absolute frequency on this tuning curve characteristic and has the additional advantage of being independent from the location of the CF within the FRA, i.e., it is not influenced by the symmetry of the FRA or tuning curve. For single units, the correlation between bandwidth and CF was statistically not significant over the limited frequency range sampled in this study.

The spatial distribution of the integrated excitatory bandwidths for pure tones obtained in this study was consistent with that of previous reports (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a). Locations with sharply tuned multiple-unit responses yielding a narrow integrated excitatory bandwidth were found in a region near the dorsoventral center of the straight segment of isofrequency contours 1.5–3 mm dorsal to the AI/AII border. Locations dorsal and ventral to this point gradually showed more broadly tuned responses.

Figure 3 illustrates multiple-unit bandwidth maps obtained 10 and 40 dB above minimum response threshold for two individual cases. *A* and *B* show the reconstructed BW40 maps. Circles connected by dotted lines show the actual bandwidth values. These and other cases show a distinct bandwidth minimum surrounded by areas with broader frequency tuning. Case SUTC16 (*B* and *D*) showed a larger variability in sharpness of tuning in areas that contained, on the average, wider bandwidths. The dorsoventral progression of bandwidth changes was much
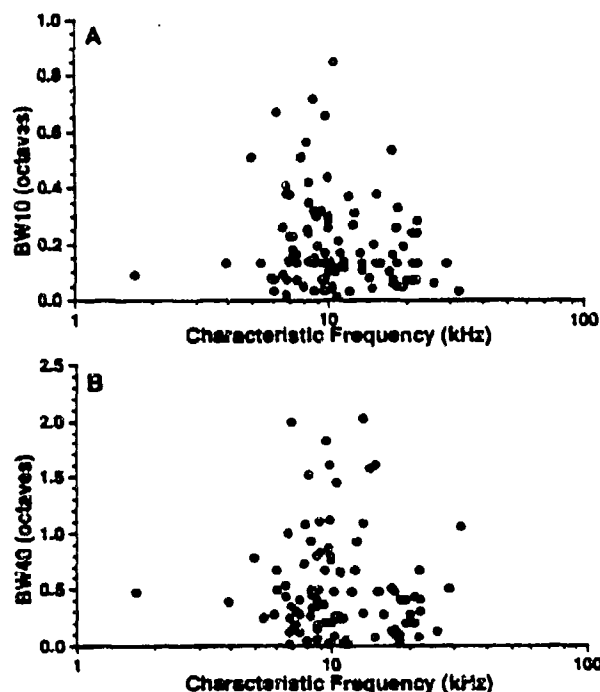

FIG. 2. Distribution of BW40 and BW10 as a function of characteristic frequency for all single neurons.
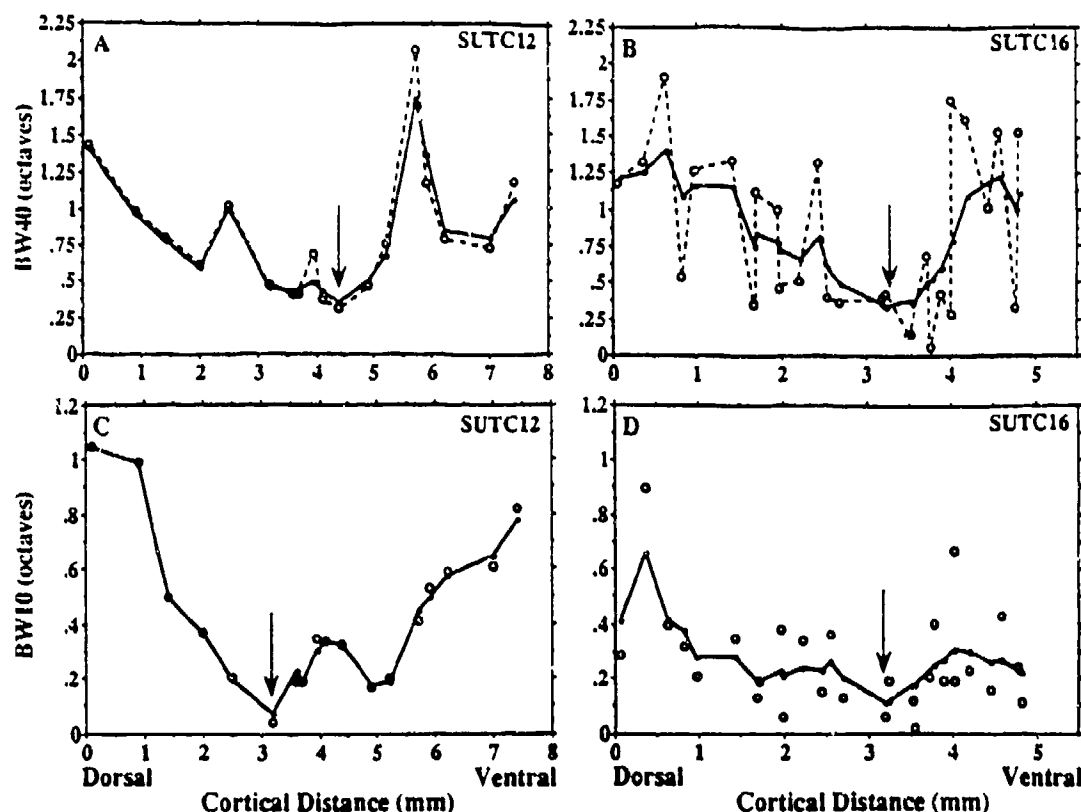
FIG. 3. Distribution of multiple-unit BW10 and BW40 values along the dorsal-ventral axis of AI for 2 cortexes. *A* and *B*: BW40 values (circles) are connected by dashed lines. Solid lines represent smoothed bandwidth distribution. A spatial weighting algorithm was used for smoothing across 500-μm sectors: 100% smoothing for points at the same location to no weighting for points 500 μm apart. Arrows mark the locations with the minimum average BW40 ($BW40_{min}$). *C* and *D*: BW10 values (circles) and smoothed bandwidth distribution. Arrows mark $BW10_{min}$.

smoother in case SUTC12 (*A* and *C*). To aid in demarcating the minima of bandwidth plots, a 500-μm, spatially weighted smoothing algorithm was used (100% smoothing for points at the same cortical distance to no weighting for points >500 μm apart). The smoothed version of the BW40 maps are shown as solid lines in Figs. 3, *A* and *B*, respectively. Arrows point to the location of the main minimum in each bandwidth function. The data points and the smoothed version of the BW10 map for these two cases are shown in Fig. 3. *C* and *D*. In both cases the BW10 minima are located 0.1–1 mm dorsal to the BW40 minimum in agreement with previous findings for Q-10dB and Q-40dB (Schreiner and Mendelson 1990). The minima of these functions ($BW10_{min}$ and $BW40_{min}$) served as physiological reference points to spatially pool excitatory bandwidth data across animals.

Single-neuron maps

The spatial distribution of excitatory bandwidths for single neurons and multiple units is shown in Fig. 4 for case SUTC16. For this animal a sufficient number ($n = 27$) of single neurons were sampled to directly compare single- and multiple-unit topographies within the same animal. The origin of the abscissa designates the location of the combined bandwidth mini num, that is, the averaged location of the BW10 and BW40 minima. For BW40 (Fig. 4*A*), the spatial distribution of the single-neuron sharpness was

quite similar to the integrative excitatory bandwidth. Neurons with the narrowest bandwidth were located at or near the multiple-unit bandwidth minimum. Dorsal to the minimum, the bandwidth for single neurons increased (linear regression analysis. $r = 0.84$, $P = 0.02$) and showed a scatter similar to that expressed in the multiple-unit responses ($r = 0.95$, $P = 0.001$). Ventrally, the excitatory bandwidth of single neurons also increased progressively ($r = 0.57$, $P = 0.02$), although not as much as seen for the distribution of multiple units ($r = 0.92$, $P = 0.001$). Sharply tuned neurons were found throughout the ventral region of AI, and, overall, the range and scatter of the integrated excitatory bandwidth appeared to be larger than for the single-neuron bandwidth. The distribution of BW10 values for single and multiple units (Fig. 4*B*) was also fairly similar. In this particular case the minimum of the bandwidth distribution was not very strongly expressed for either multiple (see Fig. 3*D*) or single neurons. However, the apparent scatter of encountered bandwidths appeared to be the smallest near the estimated center of the bandwidth distribution. The correlation between location and BW10 was statistically significant only for multiple units in the dorsal part of AI ($r = 0.81$, $P = 0.001$). The CFs of the single- and multiple-unit neurons are shown in Fig. 5. In this particular case, the CFs covered a range of ~1 octave. Although the multiple-unit distribution shows a clear correlation between dorsoventral distance and CF (dorsal half: $r = 0.65$, $P = 0.01$; ventral
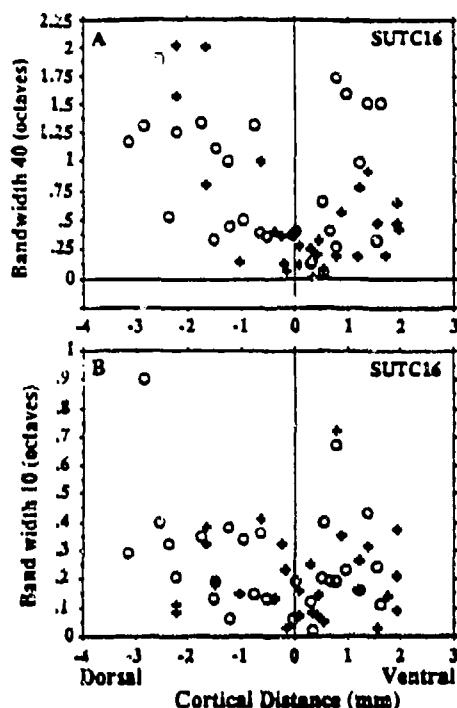
FIG. 4. Bandwidth distribution along approximated isofrequency domain of AI (case SUTC16). BW40 (*A*) and BW10 (*B*) for single-neuron (+) and multiple-unit (o) are plotted as a function of cortical distance. Zero millimeter of the cortical distance axis corresponds to the location of $BW40_{min}$.

half: $r = 0.83$, $P = 0.001$), no significant correlation between CF and location or bandwidth was seen for single neurons in either half of AI (see DISCUSSION).

Pooling relative to multiple-unit maps

The reconstruction of the spatial bandwidth distribution for case SUTC16 was made possible by sampling a relatively large number of single (27) and multiple (28) units in the same experiment. This, however, was an exceptional case, and therefore a method of pooling data across animals was necessary to generally assess the topography of single-neuron characteristics.

In this study the location of single neurons within an experiment was referenced relative to the multiple-unit bandwidth map. The alignment and pooling procedure is illustrated in Fig. 6 for the two cases shown in Figs. 3 and 4. To aid in visualizing this procedure, single neurons from case SUTC12 are marked by filled diamonds, and single neurons from case SUTC16 are marked by crosses. The unsmoothed BW40 maps for cases SUTC12 and SUTC16 (from Fig. 3. *A* and *B*) are represented by open circles connected by lines in Fig. 6, *A* and *B*. For these two experiments the dorsal-ventral extent of AI was ~6.2 and 5.3 mm, respectively. To pool data across these two experiments, the minimum in the smoothed BW40 distribution ($BW40_{min}$ marked by arrows) was assigned the reference value of 0 mm, and the positions of all recording points were expressed as distance in millimeters from $BW40_{min}$. Next, the $BW40_{min}$ of all cases was aligned. The shifted and aligned recording locations of the two cases are shown in

Fig. 6. *C* and *D*. Finally, the locations of single neurons were pooled across both aligned maps as displayed in Fig. 6 *E*. No distance normalization or distortion of the individual recording locations is involved in this method, and, accordingly, the pooled data reflect the trends discernible in each individual case. This procedure was used to pool data across all nine cases.

BW40 Topography

Scattergrams of BW40 values for all pooled single and multiple units are shown in Fig. 7. *A* and *B*. Cortical locations were aligned relative to $BW40_{min}$ for each individual case as described above. For multiple-unit responses, the pooled BW40 data show the same pattern seen in individual cases in this and previous studies (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a). In particular, a region of sharp tuning, located at the approximate dorsoventral center of AI, was aligned with integrated excitatory
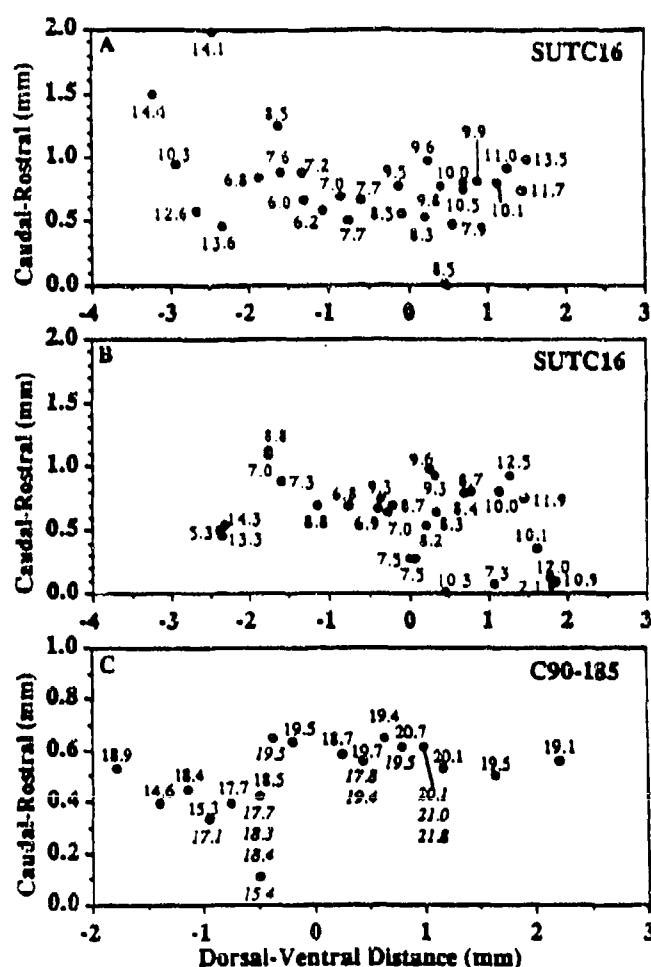
FIG. 5. Cortical locations of characteristic frequencies (CFs) for multiple and single units. *A* and *B*: locations of multiple- and single-unit recordings, respectively, for case SUTC16. CFs (in kHz) are indicated next to the location marker. *C*: both single- and multiple-unit recording locations for case C90-185. Multiple-unit CFs are indicated above the location marker, and single-unit CFs are indicated below the markers by italic numbers. At some locations, several single units were isolated. Note that the scale of the caudal-rostral distance is expanded compared with the dorsal-ventral distance.
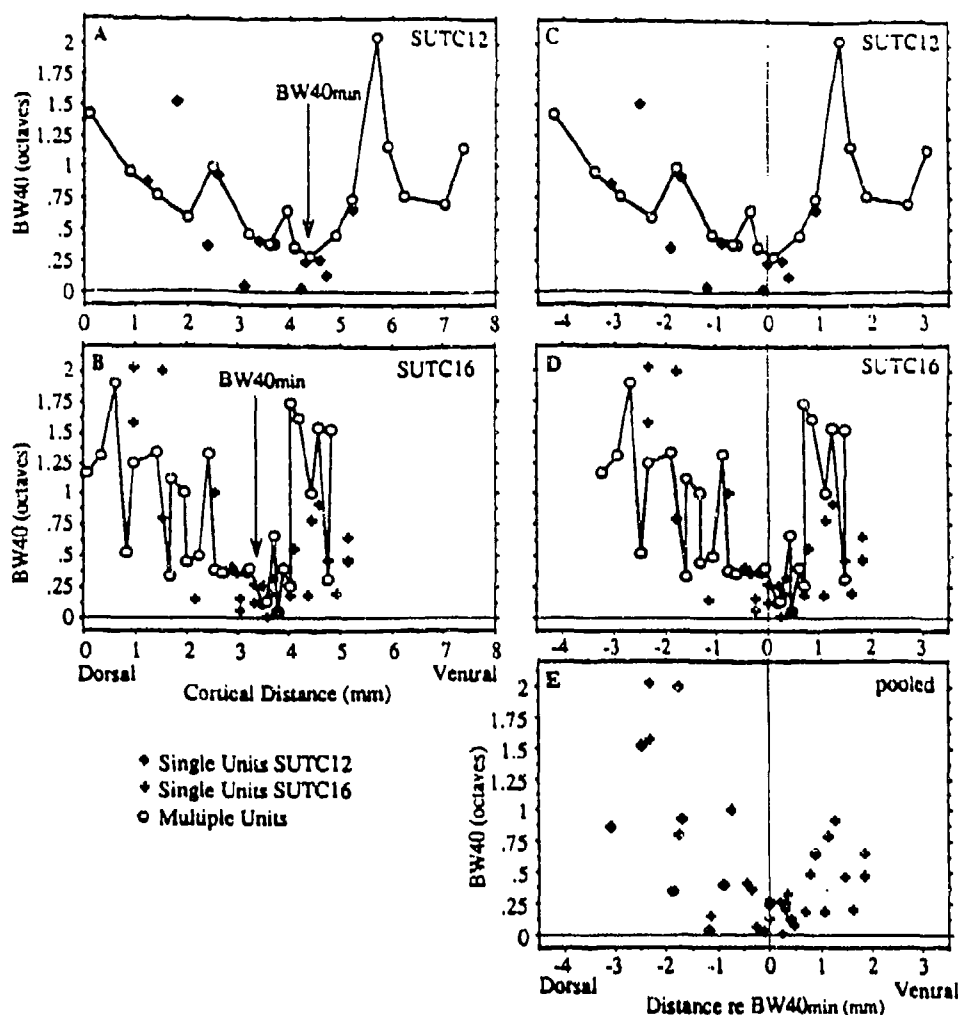
FIG. 6.   Demonstration of method used to pool data across animals. *A* and *C*: transformation for case SUTC12. Bandwidths of single neurons 40 dB above minimum threshold (diamonds) are plotted on same location axis as the multiple-unit BW40 map (connected circles). In *A*, the most dorsal location is arbitrarily assigned the value 0 mm. In *C*, $BW40_{min}$ is assigned a value of 0 mm, and all other locations are referenced as millimeters dorsal (negative) or ventral (positive) from this $BW40_{min}$. *B* and *D*: same for case SUTC16 with single-neuron BW40s depicted by crosses. $BW40_{min}$ for the 2 case is aligned (see *C* and *D*), and then the single neurons are superimposed within the same coordinate frame (*E*) to yield the pooled result.

bandwidths that gradually increased with increasing distance.

The pooled distribution of BW40 for single neurons (Fig. 7 *B*) shows some similarities and some dissimilarities to the multiple-unit distribution. Dorsal to the BW40 minimum, single units show the same tendency of increasing bandwidth with dorsal distance as seen with multiple units. Partially responsible for this overall increase in bandwidth is the occurrence of multipeaked FRAs in the dorsal aspect of AI (Fig. 7 *B*, +) in accordance with a previous report of their spatial distribution (Sutter and Schreiner 1991a). Single-peaked FRAs in dorsal AI also tended to have broader tuning. However, some sharply tuned single- and multiple-neuron responses can be found across the entire dorsal half of AI or at least up to 3 mm dorsal to the bandwidth minimum.

In the AI portion ventral to the bandwidth minimum, the pooled single-neuron data deviated from the trend seen in multiple-unit recordings as well as from that seen in the single-neuron example shown above (Fig. 4): BW40 of almost all ventral single neurons remained essentially sharply tuned (<0.7 octaves) independent of their location in ventral AI. By contrast, the multiple-unit recordings showed increasing bandwidths with distance from the BW minimum. The location of the AI/AII border varied from 1.65 to 3.45 mm from $BW40_{min}$, and no single neurons located within AII were included. However, even locations as close as 1 mm ventral to the BW40 minimum, and therefore located well within AI, showed multiple-unit locations with bandwidths in excess of 1 octave, whereas all ventral single neurons remained below that value. Figure 7*C* shows a direct comparison of the single- and multiple-unit scattergrams. BW40 values were binned and averaged over 500 μm. The mean BW40 values of multiple ($\square$) and single ($\blacksquare$) units are shown. Standard deviations are represented by error bars. The mean and standard deviation of BW40 in-
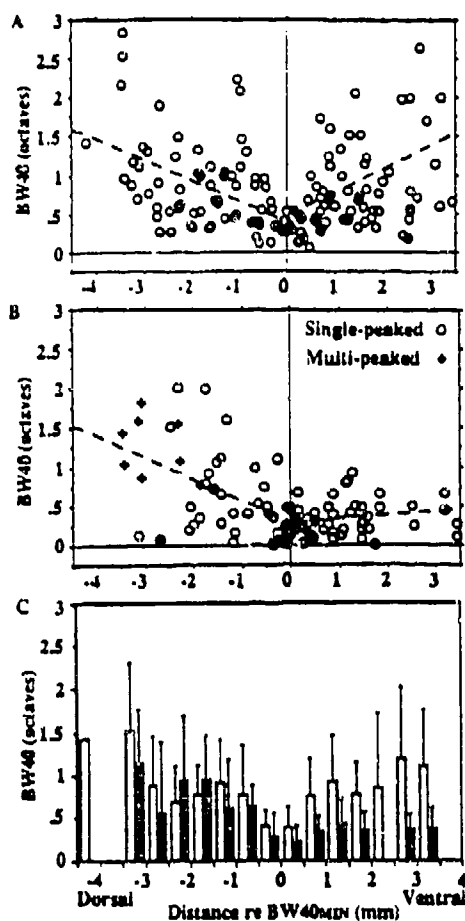
FIG. 7. Pooled BW40 data. Reference for pooling is the most sharply tuned location 40 dB above threshold of multiple-unit responses, BW40$_{min}$. *A*: distribution for all multiple-unit recordings. *B*: distribution for all single-unit recordings. *C*: mean value of BW40 over 0.5 mm bins for multiple (□) and single (■) units. Standard deviations are represented by error bars. Dashed lines in *A* and *B* correspond to the regression lines derived independently for the areas dorsal and ventral of BW$_{min}$.

creased gradually toward the dorsal end of AI for multiple and single units. Ventrally, mean and standard deviation of BW40 for single neurons remained constant, whereas the multiple-unit values increased toward and into AII.

The gradients of the bandwidth progressions dorsal and ventral to the bandwidth minimum can be approximated by regression lines as is shown by the dashed lines in Fig. 7, *A* and *B*. For the dorsal portion of the multiple-unit distribution, the regression line is described by $y = -0.25x + 0.42$ ($r = 0.49$). The slope of the dorsal gradient is -0.25 octaves/mm. The 95% confidence interval of the slope lies between -0.14 and -0.38 octaves/mm. The regression line of the ventral multiple-unit BW40 gradient was $y = 0.23x + 0.47$ ($r = 0.38$). The slope, 0.23 octaves/mm, is similar although slightly less than that of the dorsal gradient. The larger scatter on the ventral side is reflected in a 95% confidence interval containing slopes between 0.09 and 0.36 octaves/mm. For both sides of the distribution, the zero slope lies outside of the 99.5% confidence intervals. Therefore it is concluded that the multiple-unit BW40 dis-

tribution along the dorsoventral axis of AI is concave with a minimum (V shaped).

Contrasting the multiple-unit distribution, the single-unit topography (Fig. 7 *B*) appeared to be essentially flat on the ventral side. The slope of the regression line ($y = 0.05x + 0.28$; $r = 0.03$) was 0.05 octaves/mm. The 95% confidence intervals comprise slopes between -0.04 and 0.13 octaves/mm, which includes zero slope. Therefore it cannot be claimed with 95% confidence that the slope is greater than zero. Dorsally, the slope of the average excitatory bandwidth was -0.28 octaves/mm ($y = -0.28x - 0.28$; $r = 0.48$; 95% confidence interval: -0.13 to -0.42 octaves/mm), that is, similar to the dorsal slope obtained for multiple units.

As apparent from Fig. 5, case SUTC16 showed some systematic variation of CF along the dorsoventral axis of AI. This was partially due to some obscurity in the exact orientation of isofrequency contours during the experiment and to the usually encountered caudal deviation of isofrequency contours in the dorsal portion of AI. To test whether the bandwidth changes in the dorsoventral direction of AI were indeed a spatial characteristic and were not just an effect of the sampled frequency distribution along that axis, a regression analysis of BW40 with dorsoventral location as well as with CF was conducted for the dorsal and the ventral portion of all the obtained maps. The variance in the BW40 data that could be attributed to changes in the dorsoventral locations was 214 (ventral) and 318% (dorsal) larger than the variance attributable to CF changes. This suggests that the BW changes seen in the dorsoventral extent of AI are also present within isofrequency contours.

Parceling of AI: regional differences in BW40

To test the hypothesis that the dorsoventral extent of AI may consist of more than one physiologically distinguishable region, AI was subdivided into three regions on the basis of the bandwidth distribution: 1 mm to each side of the location of minimum bandwidth was arbitrarily defined as the "central" region of AI; dorsal and ventral to this area were defined correspondingly. The minimum bandwidth measure used to pool data for this analysis was the location of BW40$_{min}$ and BW10$_{min}$ averaged for each particular case (BW$_{min}$). This measure was used so that the same topographical parceling can be performed without biasing toward either bandwidth measure. Figure 8 shows the histograms of BW40 for the dorsal, ventral, and central regions of AI for single (■) and multiple (□) units. Median values for the histograms are summarized in Fig. 8C. Deviation bars represent one-half of the 75th percentile BW value minus the 25th percentile value.

Table 1 lists the mean BW40 values found with single- and multiple-unit recordings. A significant difference between the BW40 distributions of multiple and single units in the ventral and central parts of AI was established (nonparametric Mann-Whitney U test, $P < 0.01$, see Table 2). In dorsal AI, no significant difference was found ($P > 0.05$).

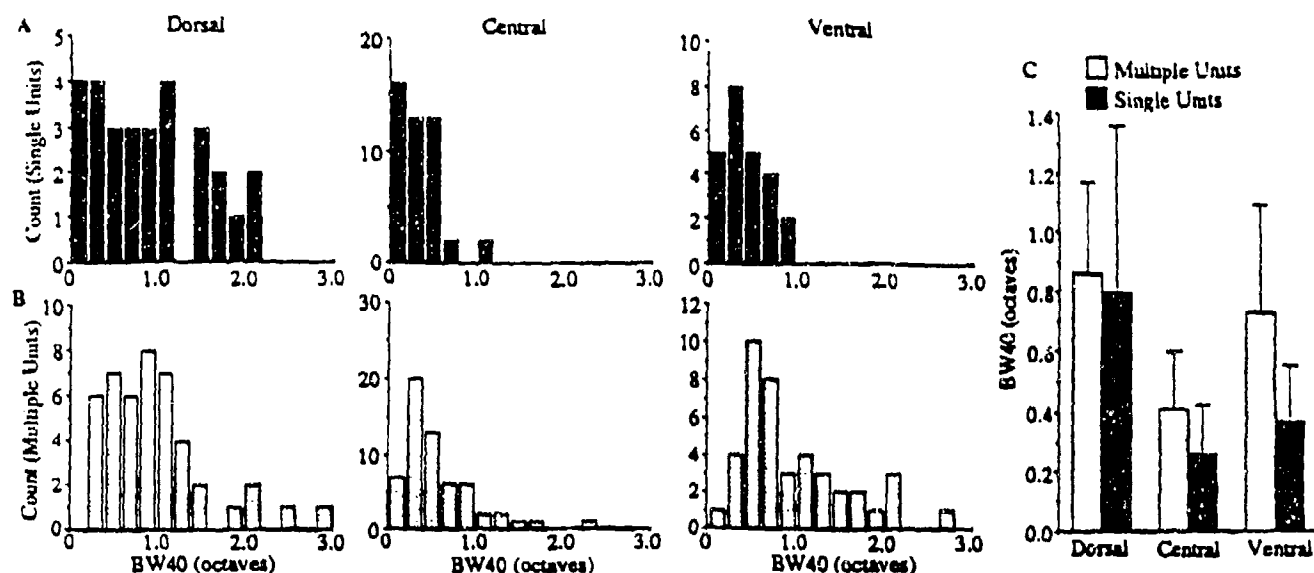Analysis of variance (ANOVA) was applied to test for differences in bandwidth values of the three regions. Signifi-

FIG. 8. Distribution of BW40 in 3 regions of the dorsoventral dimension of AI (characteristic frequency, >4 kHz). Central region includes area 1 mm to both sides of BW$_{min}$ for a total of 2 mm. Dorsal region comprises all of AI >1 mm dorsal to BW$_{min}$, and ventral region >1 mm ventral to BW$_{min}$. A: histograms of BW40 values for single neurons in AI. B: BW40 for multiple-unit recordings. C: mean values (bars) and standard deviations (error bars) of BW40 values shown in A and B.

cant differences ($P < 0.01$) between regions were found for both single- and multiple-unit BW40 values (Table 3). Post hoc Sheffé analysis showed that the ANOVA results for multiple units were due to differences between the central and ventral regions as well as between central and dorsal differences for multiple units. For single units, ANOVA differences could be accounted for by dorsal versus central and dorsal versus ventral regional differences. Nonparametric Mann-Whitney U tests confirmed the Sheffé results.

BW10 Topography

The bandwidth or Q-value 10 dB above threshold is often less appropriate for a description of sharpness of frequency tuning of single cortical neurons than, for example, Q40dB, because some influences that shape FRAs are not fully activated at stimulus levels near minimum threshold (Suga and Manabe 1982; Suga and Tsuzuki 1985). However, because this measure has been almost exclusively used in cortical studies of species other than bats and it does show some systematic spatial variations in multiple-unit recordings (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a), it was analyzed here as well to allow comparisons with other studies.

Scattergrams of BW10 for single and multiple units are shown in Fig. 9. Locations were pooled such that BW10$_{min}$ was assigned the distance reference value of zero for each case. Although the multiple-unit distribution appeared to be V shaped, paralleling the multiple-unit BW40 distribution, such an effect appeared to be less pronounced in the distribution of BW10 for single neurons. For both multiple- and single-unit recordings, FRAs narrowly tuned 10 dB above minimum threshold were encountered throughout the entire dorsoventral dimension of AI. Dorsal to the bandwidth minimum, the slope of the regression line was -0.10 octaves/mm (dashed line: $y = -0.10x + 0.14$; $r = 0.5$; 95% confidence interval: -0.06 to -1.15 octaves/mm) for multiple units. Ventrally, the slope was 0.11 octaves/mm ($y = 0.11x + 0.16$; $r = 0.37$; 95% confidence interval: 0.05–0.18 octaves/mm). For both dorsal and ventral halves, zero slope fell outside of the 99.5% confidence interval, indicating that the distribution is V shaped with 99.5% confidence and thus is similar to the BW40 distribution.

Regression analysis of single-unit BW10 topography (dashed lines in Fig. 9B) did not show slopes significantly different from zero either dorsally or ventrally. Dorsally, the slope was -0.03 octaves/mm ($y = -0.03x + 0.19$; $r = 0.14$; 95% confidence interval: -0.08–0.02 octaves/mm).

TABLE 1. *FRA bandwidth measure for single- and multiple-unit recordings in AI*

	Total	Dorsal	Central	Ventral
BW40, MU	0.80 ± 0.56 (146)	0.98 ± 0.59 (45)	0.57 ± 0.42 (59)	0.95 ± 0.59 (42)
BW40, SU	0.50 ± 0.46 (99)	0.88 ± 0.61 (29)	0.31 ± 0.24 (46)	0.40 ± 0.25 (24)
BW10, MU	0.30 ± 0.24 (146)	0.35 ± 0.23 (45)	0.22 ± 0.14 (59)	0.37 ± 0.32 (42)
BW10, SU	0.20 ± 0.17 (103)	0.26 ± 0.21 (31)	0.19 ± 0.14 (47)	0.16 ± 0.14 (25)

Values are means ± SD; number of neurons are in parentheses. Bandwidth expressed in octaves. The central region represents a 2-mm sector of the dorsoventral extent of AI centered at BW$_{min}$; the dorsal and ventral regions cover the remainder of AI. FRA, frequency response area; AI, primary auditory cortex; BW40 and BW10, bandwidth 40 and 10 dB above minimum threshold, respectively; MU, multiple units; SU, single units.

TABLE 2. *Comparison of FRA bandwidth for single- and multiple-unit measures*

	Dorsal	Central	Ventral
BW40, MU vs. SU	0.51	0.0001	0.0001
BW10, MU vs. SU	0.034	0.099	0.0001

P values from Mann-Whitney *U* tests. FRA, frequency response area: BW40 and BW10, bandwidth 40 and 10 dB above minimum threshold, respectively; MU, multiple units; SU, single units

Ventrally, the gradient was 0.01 octaves/mm ($y = 0.01x + 0.16$; $r = 0.06$; 95% confidence interval: -0.04 and 0.06 octaves/mm). It is concluded that the sharpness of frequency tuning 10 dB above minimum threshold does not change systematically for single neurons across the dorsoventral extent of AI.

Parceling of AI: regional differences in BW10

For BW10, AI was parceled as described for BW40. Histograms are shown for single (Fig. 10A, ■) and multiple (Fig. 10B, □) units in dorsal, central, and ventral AI. The summary histogram of medians is shown in Fig. 10C. The mean BW10 values for the three regions are given in Table 1. There was a significant difference between BW10 for single and multiple units in the ventral ($P < 0.01$) and dorsal ($P < 0.05$) regions of AI (see Table 2). Centrally, there was no significant difference for BW10 between single and multiple units ($P > 0.05$).

For multiple units, ANOVA revealed significant differences among dorsal, ventral, and central regions (Table 3). Sheffé post hoc tests showed the significant ANOVA was due to differences between the dorsal and central region and between the central and ventral region. Mann-Whitney *U* tests confirmed the Sheffé results (Table 2). For single units, ANOVA (confirmed by Kruskal-Wallis test) showed no significant differences in BW10 among central, ventral, and dorsal regions in AI (Table 3).

Scatter of CF in dorsoventral extent of AI

The apparent discrepancy between single- and multiple-unit estimates of the excitatory bandwidth in the ventral portion of AI suggests the influence of other physiological parameters than single-unit bandwidth on the multiple-

TABLE 3. *Comparison of FRA bandwidth measures between different sectors of AI*

	Dorsal vs. Central	Central vs. Ventral	Dorsal vs. Ventral
BW40, MU	0.0001*	0.0001*	0.64
BW40, SU	0.0001*	0.13	0.003*
BW10, MU	0.0005†	0.001*	0.96
BW10, SU	0.20	0.50	0.10

P values from Mann-Whitney *U* tests. FRA, frequency response area; AI, primary auditory cortex; BW40 and BW10, bandwidth 40 and 10 dB above minimum threshold, respectively; MU, multiple units; SU, single units. *Ninety-nine percent significance in Sheffé test. †Ninety-five percent significance in Sheffé test.
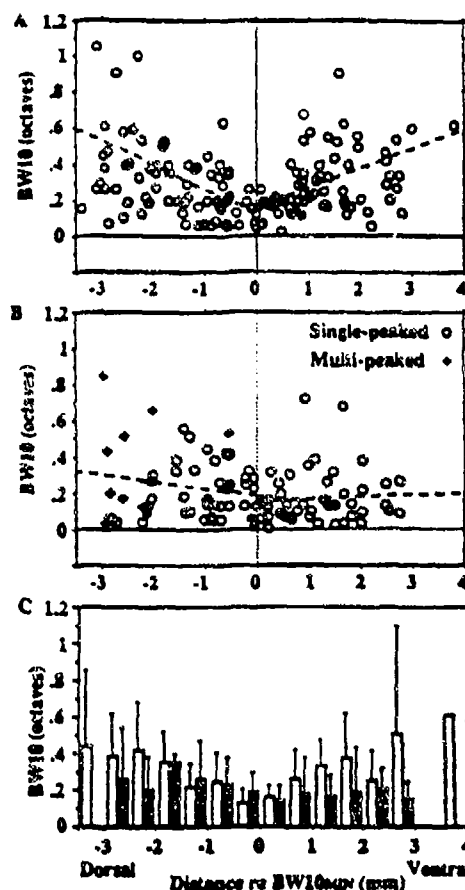
FIG. 9. Pooled BW10 data. Reference for pooling is the most sharply tuned location 10 dB above minimum threshold of multiple-unit responses, BW10$_{min}$. A: distribution for all multiple-unit recordings. B: distribution for all single-unit recordings. C: mean value of BW10 over 0.5 mm bins for multiple (□) and single (■) units. Standard deviations are represented by error bars. Dashed lines correspond to the regression lines independently derived for the areas dorsal and ventral of BW10$_{min}$.

unit result. More specifically, it indicates that, near minimum threshold of multiple-unit recordings, there are topographical differences in CF scatter and/or scatter in the minimum threshold of the contributing single neurons.

As a preliminary test of the hypothesis of a varying degree of CF scatter across AI, a post hoc analysis of single neuron CFs relative to multiple-unit CFs was conducted. The CF scatter was determined by one of two methods. If a multiple-unit CF was available for the same cortical penetration as that of a single neuron, the difference between the two CFs was determined (both single- and multiple-unit responses were obtained in the same range of cortical depth; see METHODS). For single-neuron locations that did not coincide with a penetration that yielded a multiple-unit measurement but were located between at least two multiple-unit recording locations, the multiple-unit CF at the location of the single neurons was estimated through linear interpolation/triangulation from the neighboring recording locations. The difference between the estimated multiple-unit CF and the actual single-neuron CF was obtained. Some single neurons ($n = 7$) had to be excluded from this analysis because expected CF values could not be deter-

FIG. 10. Distribution of BW10 in 3 regions of the dorsoventral dimension of AI (characteristic frequency, >4 kHz). Central region includes area 1 mm to both sides of BW$_{min}$ for a total of 2 mm. Dorsal region comprises all of AI > 1 mm dorsal to BW$_{min}$, and ventral region >1 mm ventral to BW$_{min}$. *A*: h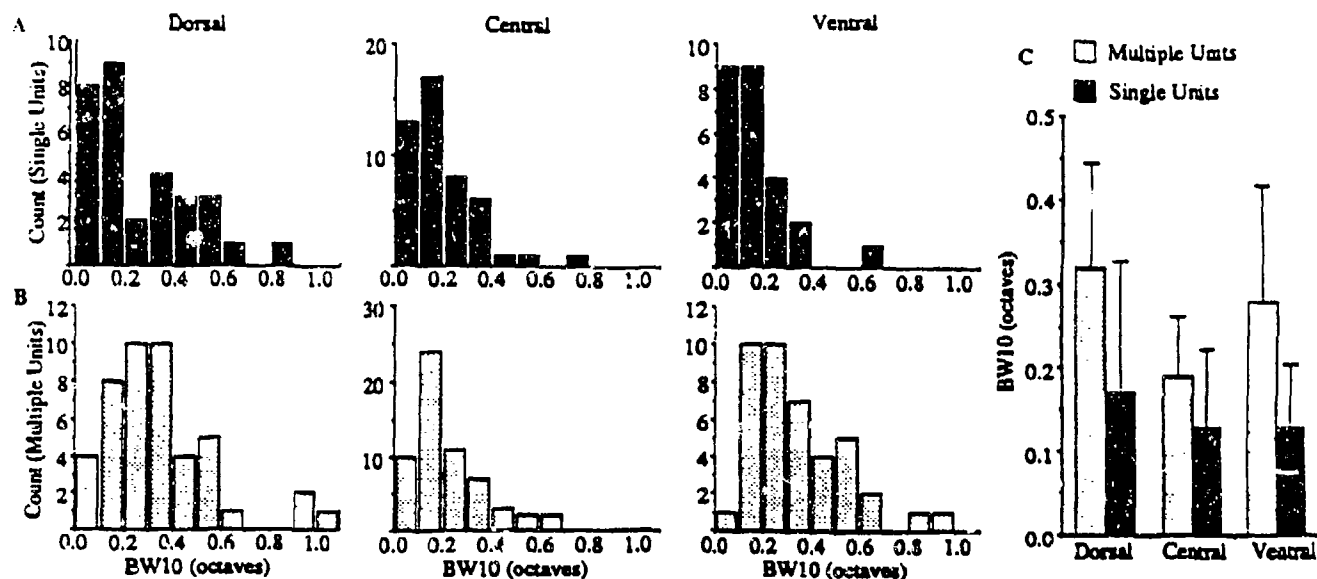istograms of BW10 values for single neurons in AI. *B*: BW10 for multiple-unit recordings. *C*: mean values (bars) and standard deviations (error bars) of BW10 values shown in *A* and *B*.

mined because of lack of appropriately positioned multiple-unit recording sites (<500-μm distance from the single-unit site). In Fig. 11, the actual and estimated CFs for case SUTC16 are plotted for multiple units (*A*) and single neu-

FIG. 11. Predicted vs. actual characteristic frequencies (CFs) for multiple and single units of case SUTC16. The predicted CFs were derived through interpolation from several multiple-unit CFs closest to a given recording site. Dashed lines indicate identity of prediction and actual CF. Formulas of the regression lines and the corresponding correlation coefficients are displayed.

rons (*B*). The dashed lines correspond to CF estimates that are identical to the actual CF. The apparent difference in CF scatter for multiple and single neurons is confirmed by a regression analysis yielding correlation coefficients of 0.89 for multiple units and 0.35 for single units. The average deviation of the actual CF from the predicted values was 0.33 ± 0.51 (SD) octaves for single units compared with 0.12 ± 0.09 octaves for multiple units ($F = 4.53$; $P < 0.04$). The spatial distribution of the CF scatter for single neurons is shown in Fig. 12. The largest local CF deviations are located at the ventral and dorsal extremes of the mapped area.

The resulting mean CF scatter, expressed in octaves, for all cases (96 single neurons) is shown in Fig. 13 as seven 1-mm-wide sectors across the dorsoventral extent of AI. Near the bandwidth minimum (at 0 mm), the obtained mean CF scatter was smallest. More dorsally and especially

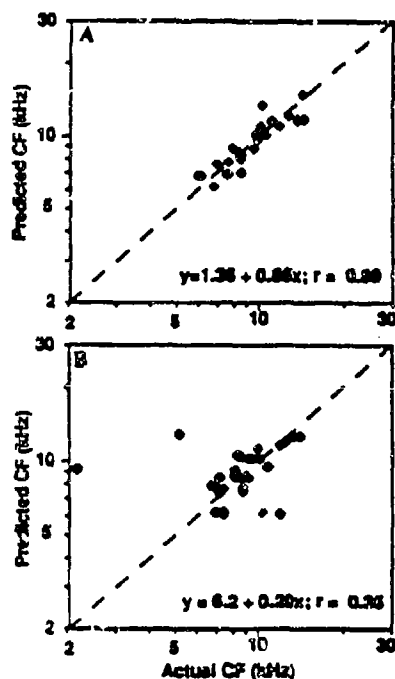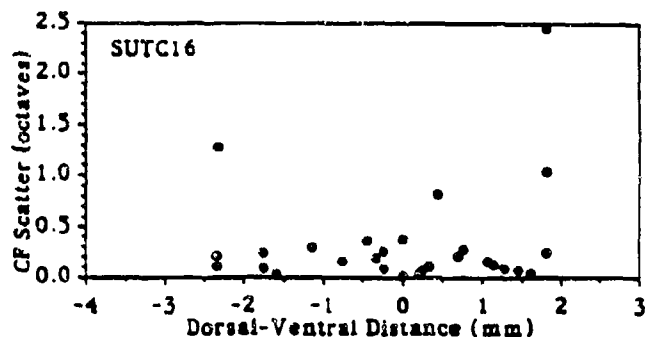

FIG. 12. Dorsoventral distribution of characteristic frequency (CF) scatter. The difference of the predicted CFs and the actual CFs of single units (SUTC16), expressed in octaves, are plotted as a function of the dorsoventral location.
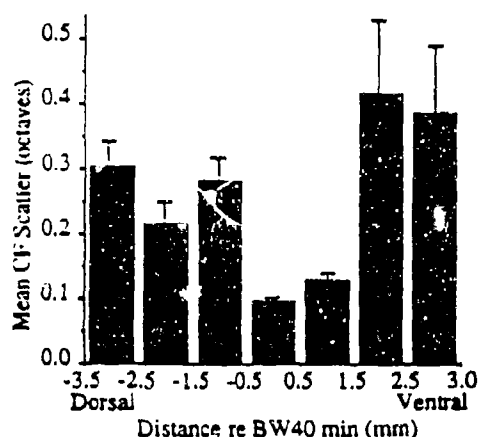
FIG. 13. Estimated scatter of characteristic frequency (CF) as a function of cortical location for 96 out of 103 single units. CF scatter = |single-neuron CF - expected CF|. Values were then averaged over 1 mm of the dorsoventral extent of AI to derive a mean CF scatter value. Number of neurons in each of the 7 bins (from dorsal to ventral) were 9, 13, 12, 28, 20, 8, and 6, respectively. Error bars indicate standard error of the mean.

more ventrally, the mean CF scatter increased substantially. The range of CF scatters at different locations along the dorsoventral axis of AI was slightly smaller than the average BW10 range for multiple units at corresponding locations (for comparison, see Fig. 9C).

DISCUSSION

In this series of experiments, the bandwidth of FRAs 10 and 40 dB above minimum response threshold was used to investigate 1) how single-neuron response properties relate to physiological "gradients" obtained with multiple-unit recording techniques, and 2) whether pooling of data across animals can be successfully used to reveal general physiological topographies. The main findings of this study can be summarized as follows. 1) Multiple-unit bandwidth data, pooled along the dorsoventral axis of AI, reflect the same spatial distribution as previously seen in individual cases (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a), namely, a region ~1.5–3 mm dorsal to the AII border contained narrow integrated excitatory bandwidths gradually giving way to more broadly tuned responses dorsally as well as ventrally. 2) Excitatory bandwidths of single neurons, pooled relative to the multiple-unit bandwidth topography, significantly differed from the multiple-unit distribution in the region of AI ventral to the bandwidth minimum, that is, no clear gradient in sharpness of single-neuron responses was evident for either BW10 or BW40; in the region of AI dorsal to the minimum in multiple-unit bandwidth, a spatial gradient as well as an increasing variance was evident for the BW40 of single neurons similar to the one obtained for multiple units, however, no clear gradient was apparent for BW10 of single neurons. 3) Predictions of the CF of a given cortical location from the CF of neighboring multiple-unit CFs resulted in a good agreement with the actual multiple-unit CFs. By contrast, a substantial deviation of single-unit CFs from multiple-unit estimates was seen for single neurons located near the ventral and dorsal ends of AI.

Methodological considerations

Before discussing some implications of these findings for the interpretation of the functional organization of primary auditory cortex, a brief look at the pooling method is in order. Single- and multiple-unit properties were pooled along the dorsoventral extent of AI, approximating the orientation of the isofrequency axis of AI. The location of the minimum in multiple-unit bandwidth either 10 or 40 dB above minimum response threshold was used as a reference point for the pooling of data, allowing an alignment of isofrequency contours from different cortexes. The use of a single point of reference plus the orientation of the contours does not involve normalization of the length of the dorsoventral extent of AI. Reale and Imig (1980) gave estimates of the dorsoventral extent of AI ranging from 4 to 7 mm. However, precise anatomic or physiological criteria for establishing the extent of that dimension are not available (see Middlebrooks and Zook 1983; Schreiner and Cynader 1984; Schreiner and Mendelson 1990; Sutter and Schreiner 1991a; Winer 1984). Therefore no attempts at extent normalization have been made. A potential consequence of differences in the absolute extent of spatial gradients may be a blurring of the topographies at the dorsal and ventral extremes of the pooled data (see below).

The pooling method did not take into account the actual CF of the units. The question is, then, could CF sampling biases have strongly affected the observed spatial bandwidth distributions? It has been reported that, with increasing CF, Q-factors of cortical neurons show an overall increase and relative bandwidth values decrease (Phillips and Irvine 1981). Previous studies (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a) had shown a similar dorsoventral pattern of tuning sharpness across all frequencies, at least for the CF range >4 kHz, which is the same range covered in the current study. In some of the cases included in this study, in particular case SUTC16, CFs at locations near the ventral and dorsal end of AI tended to deviate from the projected isofrequency values toward higher frequencies (see Fig. 5). These were also the regions that showed the widest integrative excitatory bandwidth. Because neurons with higher CFs also tend to have narrower bandwidths, a potential CF-influence would tend to diminish the magnitude of the observed bandwidth differences between the dorsoventral center and the dorsal and ventral margins of AI providing an argument against a substantial CF influence on the presented data. In addition, comparison of the variance in the bandwidth values that could be attributed to either the dorsoventral location of AI or the CF also suggested a relatively small systematic influence of CF distribution on the results. This supports the conclusion that the observed global excitatory bandwidth variations can also be found along more strictly defined isofrequency contours.

Finally, in interpreting this data, it is important to remember that the single and multiple neuron recordings from these experiments were limited to depths between 600 and 1,000 μm below the cortical surface. Therefore the presented spatial distribution of single-neuron properties may be limited to a portion of AI columns, approximately corresponding to deep layer 3 and layer 4.

Single-unit versus multiple-unit bandwidth

The multiple-unit measurements of sharpness of tuning (excitatory bandwidth 10 and 40 dB above minimum threshold) confirm the spatial pattern in the dorsoventral dimension of AI as reported for corresponding Q-values (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a). All isofrequency contours contain a region of narrow integrated excitatory bandwidths with a gradual transition to more broadly tuned locations toward the ventral and dorsal ends of AI. In other words, the excitatory contributions that are available at different cortical locations along the isofrequency domain appear to vary systematically. What is the source of this gradation in the integrated excitatory bandwidth? Pooled single-neuron data dorsal to the minimum in integrated bandwidth 40 dB above threshold showed units with progressively broader bandwidths intermingled with sharply tuned neurons. Local integration of activity from several neurons can account for the observed dorsal bandwidth gradient as reflected in similar slopes of the linear regression for multiple- and single-unit BW40 (Fig. 7). The occurrence of neurons with multipeaked tuning curves in dorsal AI (Sutter and Schreiner 1991a) contributed to the broadening of the multiple-unit BW40. However, neurons with a single excitatory response area also broadened, occasionally exceeding the total bandwidth of multipeaked neurons (Fig. 7). Broad single-peaked neurons did not necessarily have to reside in close proximity to recorded multipeaked neurons.

Ventral to the minimum in integrated bandwidth, pooled BW40 values of single neurons did not reflect the same trend as seen for the integrated bandwidth. All single neurons in the ventral 3.5 mm of AI remained essentially sharply tuned, that is, below a bandwidth of 1 octave. By contrast, ~25% of the multiple units in that area had bandwidths >1 octave. Accordingly, no bandwidth gradient was observed for the pooled single-neuron data. Some aspects of the pooling method may have contributed to this discrepancy between multiple- and single-unit measurements. Single units that were located in AII, as defined by multiple-unit criteria (Schreiner and Cynader 1984), were not included in the pooled data. A strict definition of the AI/AII border may have biased the single-unit distribution toward narrower bandwidths at the ventral end of AI. The distance between the bandwidth minimum and the border of AII could vary over >2 mm, even within a single case. Because the length of AI was not normalized for the pooling, a functional gradient between BW40 minimum and AII could be obscured because of the averaging. Indeed, in one of the illustrated single-unit cases (see Fig. 4), a BW40 gradient is apparent in ventral AI. However, the single-unit gradient appears to be somewhat shallower than the multiple-unit gradient. Additionally, even at ventral distances as close as 1 mm from the bandwidth minimum, that is well within AI, the bandwidth discrepancy between single- and multiple-unit recordings was apparent. There are reasons to believe, then, that the lack of congruence in BW40 for single and multiple units in ventral AI is not solely caused by potential methodological influences but is a reflection of physiological properties that distinguish ventral from dorsal AI (see below).

Discrepancies between single- and multiple-unit responses were more pronounced for the measure of BW10. Although the multiple-unit BW10 data showed similar gradients along the dorsoventral extent as seen for BW40 in this study and for corresponding Q-values in previous studies (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a), pooled BW10 values for single neurons showed no statistically significant bandwidth gradients in either ventral or dorsal AI. The difference between single and multiple units was most strongly expressed in ventral AI with constant average bandwidths across the most ventral 3 mm of AI. In the region dorsal to the multiple-unit BW10 minimum, single-neuron BW10 values did show a trend similar to the multiple-unit data (see Fig. 9C) and paralleling the BW40 results, however, without reaching statistical significance. Contributing to the single/multiple-unit discrepancy for BW10 in the dorsal region of AI was the fact that the different peaks in multipeaked tuning curves show differences in the minimum threshold of each peak (Sutter and Schreiner 1991a). Consequently, some of those units are represented by the bandwidth of only one peak and not the total bandwidth across all peaks, because minimum threshold of those additional peaks have not been reached 10 dB above the minimum threshold of the most sensitive peak.

It is concluded that sharply tuned neurons can be found across the entire dorsoventral extent of AI. The spatial variation in integrated bandwidth in the dorsal region of AI is paralleled by an increase in the scatter of single-unit bandwidth and by an increasing occurrence of multipeaked neurons toward the dorsal end of AI. By contrast, the increase of integrated bandwidth toward the AII border is not necessarily paralleled by an increase in single-unit bandwidth. In the approximate dorsoventral center of AI is a region of sharply tuned single- as well as multiple-unit responses.

Influence of CF scatter on bandwidth distribution

The apparent discrepancy between single- and multiple-unit estimates of the excitatory bandwidth in the ventral portion of AI was suggestive of the presence of a change in the local scatter of CFs especially from central AI to the ventral boundary. The physiological dorsoventral center of AI, defined by a narrow multiple-unit bandwidth, can only show little CF scatter of single neurons because single- and multiple-unit responses are equally sharply tuned. An increasing scatter in CF toward the dorsal, and especially ventral, borders of AI could account for the observed increase in multiple-unit bandwidth.

Although a proper estimate of the distribution of single-unit CF scatter in AI should be done by analyzing several single neurons at each sampled cortical location (e.g., Hui et al. 1989), the currently applied method of comparing multiple-unit CFs with single-neuron CFs should provide a reasonable approximation. The main justifications are that *1*) multiple-unit CFs are likely close to the average of the contributing single-neuron CFs, and *2*) multiple-unit CF distributions across AI show a very high degree of cochleotopicity (e.g., Merzenich et al. 1975; Reale and Imig 1980; Schreiner and Mendelson 1990). Although isofrequency contours (as determined with the multiple-unit method)

are not everywhere as straight as in the central sector of AI. Influences of systematic curvatures in the orientation of iso-frequency contours (e.g., Sutter and Schreiner 1991a) at the dorsal and/or ventral ends on the CF scatter estimate are likely to be small, because the estimates are either from the same locations as the multiple-unit CFs (i.e., are independent from the orientation) or are derived from several next neighbors, thus taking potential local changes in orientation into account. The current attempt to analyze the local consistency in the CF of cortical neurons indicates then that the hypothesized spatial change in CF scatter of single neurons is likely to be true. Combining the range of CF scatter with the range of single-neuron bandwidth for each AI sector results in a good approximation of the overall excitatory bandwidth distribution obtained with multiple-unit recording.

How broad multiple-unit responses can emerge from narrow single-neuron responses can be viewed by considering multiple-unit responses as the total integrated excitatory response of the entire recording site (Schreiner and Mendelson 1990). The multiple-unit recording could reflect all inputs, inhibitory and excitatory, because this method can potentially record from somas, dendrites, and axons of different neurons. With this in mind, the single/multiple-unit differences in bandwidth encountered in ventral AI might be due to the inclusion of spikes from proximal inhibitory processes in the vicinity of a postsynaptic neuron into the multiple-unit response. Therefore broad excitatory responses might be due to broad inhibitory inputs of narrowly tuned ventral neurons. If all inputs to a given cortical location, producing either inhibition or excitation, were reflected in multiple-unit recordings, dorsally located multipeaked single neurons should be accompanied by broad single-peaked multiple-unit recordings at the same location, which, indeed, was often the case (personal observation).

Parceling of AI

Dividing AI into three different regions provided a direct method of addressing questions regarding local differences between single- and multiple-unit bandwidth measures. The selection of three areas was based on the distribution of the integrated bandwidth alone. The combined results from single and multiple units suggest that, physiologically, AI can be parceled into two regions: a ventral (AIv) and a dorsal (AId) region. The bases of dividing these regions are *1*) a reversal of the global gradient of integrated bandwidth along the dorsoventral axis of AI, *2*) regional differences in the distribution of BW40 for single neurons, and *3*) the observation of neurons with multiple-peaked tuning curves that appear to be limited to the dorsal part of AI.

A significant difference in the spectral integration properties of the granular and supragranular layers of AId and AIv implies that there should be fundamental differences in the spectral processing in the two regions. Functionally, AId may be particularly well-suited for an integrative analysis of broadband stimuli by responding relatively undifferentiated to tones or spectral peaks of different frequencies. This is reflected in generally broader excitatory tuning of single neurons as well as by multipeaked neurons. Accord-

ingly, a high responsiveness to broadband stimuli was observed in AId (Schreiner and Mendelson 1990: Sutter and Schreiner 1991a). However, whether a true integrative processing, e.g., characterized by lower thresholds for broadband signals compared with those for tones, is present in AId remains to be seen.

Locations in AIv may be more suited for a differential analysis of the spectral properties of broadband stimuli because they are characterized by sharply tuned single neurons distributed over varying ranges of CFs. In other words, details of the spectral distribution of broadband signals appear to be resolved and preserved at given locations in AIv. The responsiveness to broadband stimuli (clicks or white noise) is generally smaller than in AId (Schreiner and Mendelson 1990), because of the relatively sharply tuned single neurons. The distribution of other response parameters, in particular minimum threshold, monotonicity of rate/level functions, and binaurality, have to be taken into account to derive a more complete interpretation of the implications of the described functional topography of AI.

Relation to previous studies

The results of this paper may help to resolve discrepancies among various reports of CF topography and sharpness of tuning in AI obtained with single- or multiple-unit measures. A highly systematic CF topography, the cochleotopic organization, has been reliably found in cat AI when applying the multiple-unit recording technique (e.g., Merzenich et al. 1975: Reale and Imig 1980: Schreiner and Mendelson 1990). By contrast, several attempts to establish the cochleotopic organization of AI with single-unit recordings have indicated limitations on the precision of such an organization on the single-neuron level (Abeles and Goldstein 1970a,b: Bogdanski and Galambos 1960: Erulkar et al. 1956: Evans and Whitfield 1964: Evans et al. 1965: Goldstein et al. 1968, 1970). Investigators who reported a lack of CF consistency also report a wide range of tuning sharpness (50% >0.5 octaves) and a relatively high percentage of multipeaked neurons. Other studies have claimed that almost all units in AI were sharply tuned (e.g., Phillips and Irvine 1981). The recording of predominantly sharply tuned neurons was possibly due to recording only from the ventral region of AI and/or a reliance on Q-10dB as the measure of sharpness.

Some of the inconsistencies between multiple-unit and earlier single-unit mapping studies may be accounted for by differences in the extent of the considered cortical area, spatial scatter induced by pooling across animals, and by the state of anesthesia (Merzenich et al. 1975). However, some discrepancies remain such as the range of CF scatter found in single, near-radial penetrations of AI (Abeles and Goldstein 1970a,b: Evans and Whitfield 1964: Goldstein et al. 1968: Phillips and Irvine 1981). The cochleotopicity obtained with the multiple-unit technique is derived from the minimum threshold, that is from the most sensitive neuron or from the average CF of the most sensitive neurons within each cluster. For single-neuron studies, an apparent lack of CF topography or even CF consistency within the same penetration (Evans and Whitfield 1964: Goldstein et al. 1968, 1970: Katsuki et al. 1959a,b) could be due to record-

ings from neurons whose thresholds and CFs are different from the most sensitive neuron(s) near the recording location. The observations in the current study indicate that the precision in the spatial CF distribution of single neurons changes gradually across AI from a strict local alignment in the central 1–2 mm of its dorsoventral extent to a decreasingly strict alignment toward the dorsal and ventral boundaries of AI. This loss of precision in single-unit CF organization is present in areas that still show a high degree of cochleotopic organization when tested with multiple-unit recordings and, therefore, has to be distinguished from the scatter of multiple-unit CFs that is encountered in the dorsal portion of AII (Schreiner and Cynader 1984).

These findings suggest that the integrated multiple-unit response is reflecting the mode of the CF distribution at a given location. Single-unit studies usually provide only a single sample per location that is subject to the variance in the composition of the local cell assembly. Accordingly, the notion of a strictly defined isorepresentational frequency domain, i.e., the pure isofrequency contour, can only be maintained for integrative response measures or for a narrow 1- to 2-mm-wide portion at the dorsoventral center of AI. The suggestion of a changing variance in the CF distribution of single neurons along the dorsoventral axis of AI consolidates the picture of the precision in frequency organization as obtained with these different recording methods. Additionally, it provides some of the physiological basis for the observation of changing integrated bandwidths in the cortical dimension orthogonal to the cochleotopic gradient.

Recent studies in primates and birds have indicated that the systematic spatial distribution of integrative excitatory bandwidth, as described earlier (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a) and in this report, might be a general organizational property of the primary auditory telencephalic field. The integrated excitatory bandwidth changes systematically along the isofrequency domain in AI of the owl monkey (Merzenich et al. 1991) in a similar fashion as seen in the cat. In a combined multiple- and single-unit study of the neostriatal field L in the chick, sharp tuning was found in the approximate center of the isofrequency domain (Heil and Scheich 1991). As in the two mammalian species, an increase in breadth of tuning gradually occurs with recording sites further away from that point.

The pooling across animals in these studies does not allow to distinguish unequivocally whether the observed global dorsoventral gradients in excitatory bandwidth indeed reflect a relatively smooth gradient in individual cases (due to increases in scatter of CF or bandwidth of the underlying single neurons) or are the consequence of a more patchy or "columnar" organization of clustered neurons with similar bandwidth characteristics. The latter hypothesis is suggested by the patchy distribution of several anatomic and functional aspects of AI organization (Imig and Adrian 1977; Matsubara and Phillips 1988; Phillips et al. 1985; Rajan et al. 1990; Wallace et al. 1991). Signs of a certain functional patchiness is also apparent in individual cases of the maps obtained in this series for spectral as well as intensity parameters (see Schreiner and Mendelson 1990; Schreiner et al. 1992). However, the strictness of

such a hypothetical columnar bandwidth organization would have to be fairly limited because any combination of single-peaked broad, single-peaked narrow, and multi-peaked broad neurons can be encountered in a given penetrations in the dorsal half of AI. The notions of 1) general trends in the functional organization of the dorsoventral extent of AI as expressed, for example, in the distribution of excitatory bandwidth and tonotopicity and 2) the occurrence of cortical patches with similar functional attributes are not necessarily incompatible. It may be argued that general spatial trends exist in the organization of primary auditory cortex and that these trends are modulated by small, functionally more coherent patches. How the physiological distinctions of neurons along the main gradients and within different cortical patches can be explained within functional, morphological, and/or projectional frameworks of cortical processing remains to be seen.

In conclusion, this study reports a physiological framework of the dorsoventral extent of AI that is based on the spatial distribution of spectral parameters, namely the excitatory bandwidth (sharpness of tuning) and the characteristic frequency of single as well as local groups of neurons in AI. The presented evidence is compatible with the notion that AI can be divided into at least two functionally separable regions: dorsal AI (AId) and ventral AI (AIv). Functional interpretations of the role of AI from physiological properties of single neurons need to take these topographically based representational principles into account.

REFERENCES

ABELES, M. AND GOLDSTEIN, M. H., JR. Functional architecture in cat primary auditory cortex: columnar organization and organization according to depth. *J. Neurophysiol.* 33: 172–187, 1970a.

ABELES, M. AND GOLDSTEIN, M. H., JR. Responses of single units in the primary auditory cortex of the cat to tones and tone pairs. *Brain Res.* 42: 337–352, 1970b.

BOGDANSKI, D. V. AND GALAMBOS, R. Studies of the auditory system with implanted electrodes. In: *Neural Mechanisms of the Auditory and Vestibular Systems*, edited by G. L. Rasmussen and W. F. Windle. Springfield, IL: Thomas, 1960, p. 137–151.

ERULKAR, S. D., ROSE, J. E., AND DAVIES, P. W. Single unit activity in the auditory cortex of the cat. *Johns Hopkins Hosp. Bull.* 39: 55–86, 1956.

EVANS, E. F., ROSS, H. F., AND WHITFIELD, I. C. The spatial distribution of unit characteristic frequency in the primary auditory cortex of the cat. *J. Physiol. Lond.* 179: 238–247, 1965.

EVANS, E. F. AND WHITFIELD, I. C. Classification of unit responses in the auditory cortex of the unanesthetized and unrestrained cat. *J. Physiol. Lond.* 171: 476–493, 1964.

GOLDSTEIN, M. H., JR., ABELES, M., DALY, R. L., AND McINTOSH, J. Functional architecture in cat primary auditory cortex: tonotopic organization. *J. Neurophysiol.* 33: 188–197, 1970.

GOLDSTEIN, M. H., JR., HALL, J. L., II, AND BUTTERFIELD, B. O. Single-unit activity in the primary auditory cortex of unanesthetized cats. *J. Acoust. Soc. Am.* 43: 444–455, 1968.

HEIL, P. AND SCHEICH, H. Functional organization of the avian auditory cortex analogue. I. Topographic representation of isointensity band-width. *Brain Res.* 539: 110–120, 1991.

HIND, J. E. Unit activity in the auditory cortex. In: *Neural Mechanisms of the Auditory and Vestibular Systems*, edited by G. L. Rasmussen and W. F. Windle. Springfield, IL: Thomas, 1960, p. 201–211.

HUI, G. K., CASSADY, J. M., AND WEINBERGER, N. M. Response properties of single neurons within clusters in inferior colliculus and auditory cortex. *Soc. Neurosci. Abstr.* 15: 746, 1989.

IMIG, T. J. AND ADRIAN, H. O. Binaural columns in the primary field (AI) of auditory cortex. *Brain Res.* 138: 241–257, 1977.

IMIG, T. J. AND BRUGGE, J. F. Relationship between binaural interaction columns and commissural connections of the primary auditory field (AI) in the cat. *J. Comp. Neurol.* 182: 637–660, 1978.

IMIG, T. J., IRONS, W. A., AND SAMSON, F. R. Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J. Neurophysiol.* 63: 1448–1466, 1990.

IMIG, T. J. AND REALE, R. Ipsilateral cortico-cortical projections related to binaural columns in cat primary auditory cortex. *J. Comp. Neurol.* 203: 1–14, 1981.

KATSUKI, Y., WATANABE, T., AND MARUYAMA, N. Activity of auditory neurons in upper levels of brain of cat. *J. Neurophysiol.* 22: 343–359, 1959a.

KATSUKI, Y., WATANABE, T., AND SUGA, N. Interaction of auditory neurons in response to two sound stimuli in cat. *J. Neurophysiol.* 22: 603–623, 1959b.

MATSUBARA, J. A. AND PHILLIPS, D. P. Intracortical connections and their physiological correlates in the primary auditory cortex (AI) of the cat. *J. Comp. Neurol.* 268: 38–48, 1988.

MENDELSON, J., SCHREINER, C. E., GRASSE, K., AND SUTTER, M. Spatial distribution of responses to FM sweeps in cat primary auditory cortex. *Assoc. Res. Otolaryngol. Abstr.* 11: 36, 1988.

MERZENICH, M. M., KNIGHT, P., AND ROTH, G. L. Representation of the cochlea within the primary auditory cortex in the cat. *J. Neurophysiol.* 38: 231–249, 1975.

MERZENICH, M. M., SCHREINER, C. E., RECANZONE, G. H., BEITEL, R. E., AND SUTTER, M. L. Topographic organization of cortical field AI in the owl monkey (*aotus trivirgatus*). *Assoc. Res. Otolaryngol. Abstr.* 14: 44, 1991.

MIDDLEBROOKS, J. C., DYKES, R. W., AND MERZENICH, M. M. Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. *Brain Res.* 181: 31–48, 1980.

MIDDLEBROOKS, J. C. AND ZOOK, J. M. Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *J. Neurosci.* 3, 203–224, 1983.

PHILLIPS, D. P. AND IRVINE, D. R. F. Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. *J. Neurophysiol.* 45: 48–58, 1981.

PHILLIPS, D. P., ORMAN, S. S., MUSICANT, A. D., AND WILSON, G. F. Neurons in the cat's primary auditory cortex distinguished by their response to tones and wide-spectrum noise. *Hear. Res.* 18: 73–86, 1985.

RAJAN, R., AITKIN, L. M., AND IRVINE, D. R. F. Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J. Neurophysiol.* 64: 888–902, 1990.

REALE, R. A. AND IMIG, T. J. Tonotopic organization in cat auditory cortex. *J. Comp. Neurol.* 192: 265–292, 1980.

REALE, R. A. AND KETTNER, R. E. Topography of binaural organization in primary auditory cortex of the cat: effects of changing interaural intensity. *J. Neurophysiol.* 56: 663–682, 1986.

ROSE, J. E. The cellular structure of the auditory region of the cat. *J. Comp. Neurol.* 91: 409–440, 1949.

SCHREINER, C. E. AND CYNADER, M. S. Basic functional organization of second auditory cortical field (AII) of the cat. *J. Neurophysiol.* 51: 1284–1305, 1984.

SCHREINER, C. E. AND MENDELSON, J. R. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophysiol.* 64: 1442–1459, 1990.

SCHREINER, C. E., MENDELSON, J. R., GRASSE, K., AND SUTTER, M. Spatial distribution of basic response properties in cat primary auditory cortex. *Assoc. Res. Otolaryngol. Abstr.* 11: 36, 1988.

SCHREINER, C. E., MENDELSON, J. R., AND SUTTER, M. L. Functional topography of cat primary auditory cortex: representation of tone intensity. *Exp. Brain Res.* In press.

SHAMMA, S. A. AND FLESHMAN, J. W. Spectral orientation columns in the primary auditory cortex. *Assoc. Res. Otolaryngol. Abstr.* 13: 222, 1990.

SUGA, N. AND MANABE, T. Neural basis of amplitude-spectrum representation in the auditory cortex of the mustached bat. *J. Neurophysiol.* 47: 225–255, 1982.

SUGA, N. AND TSUZUKI, K. Inhibition and level-tolerant frequency tuning in the auditory cortex of the mustached bat. *J. Neurophysiol.* 53: 1109–1145, 1985.

SUTTER, M. AND SCHREINER, C. E. Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65: 1207–1226, 1991a.

SUTTER, M. AND SCHREINER, C. E. Spatial distribution of the excitatory bandwidth of single neurons in cat primary auditory cortex. *Assoc. Res. Otolaryngol. Abstr.* 13: 21, 1991b.

WALLACE, M. N., KITZES, L. M., AND JONES, E. G. Intrinsic inter- and intralaminar connections and their relationship to the tonotopic map in cat primary auditory cortex. *Exp. Brain Res.* 86: 527–544, 1991.

WINER, J. A. Anatomy of layer IV in cat primary auditory cortex (AI). *J. Comp. Neurol.* 224: 535–567, 1984.

Functional topography of cat primary auditory cortex: representation of tone intensity

Christoph E. Schreiner, Julie R. Mendelson, and Mitchell L. Sutter

Coleman Laboratory, Department of Otolaryngology, W.M. Keck Center for Integrative Neuroscience University of California at San Francisco, San Francisco, CA 94143-0732, USA

Summary. The neuronal response to tones as a function of intensity was topographically studied with multiple-unit recordings in the primary auditory cortex (AI) of barbiturate-anesthetized cats. The spatial distribution of the characteristics of rate/level functions was determined in each of three intensely studied cases and their relationship to the distribution of spectral parameters (sharpness of tuning and responses to broadband transients) in the same animals was determined. The growth of the high-intensity portion of rate/level functions was estimated by linear regression. Locations with monotonically growing high-intensity portions were spatially segregated from locations with nonmonotonic rate/level functions. Two noncontiguous areas with a high degree of non-monotonicity were observed. One was located at the dorsoventral center of AI, and a second in the dorsal third of AI. The more ventral aggregate of high non-monotonicity coincided with the region of sharp frequency tuning. The stimulus levels that produced the highest firing rate (strongest response level, SRL) at any sampled location ranged from 10 to 80 dB sound pressure level (SPL). Several spatial aggregates with either high or low SRLs were observed in AI. The region of sharpest tuning was always associated with a region of low SRLs. The response threshold to contralateral tones at the characteristic frequency (CF) ranged from −10 dB SPL to 85 dB SPL with the majority between 0 and 40 dB SPL. The spatial distribution of response thresholds indicated several segregated areas containing clusters with either higher or lower response thresholds. The correlation of response threshold with integrated bandwidth and transient responses was only weak. Low- and high-intensity tones of the same frequency are represented at different locations in AI as judged by the amount of evoked neuronal activity and are largely independent of the frequency organization. The spatial distribution of locations with high monotonicity and low strongest response levels were aligned with the organization of the integrated excitatory bandwidth and covaried with the response strength to broadband stimuli.

Introduction

Studies of the spatial distribution of several functional response properties in mammalian primary auditory cortex (AI) have revealed systematic physiological organizations. The most prominent, and best documented, organizational features are a cochleotopic arrangement of the characteristic frequency (CF) of neurons (Merzenich et al. 1975; Reale and Imig 1980; Woolsey and Walzl 1942) and the spatial segregation of different binaural integrations of excitatory and inhibitory inputs into "binaural interaction bands" (Imig and Adrian 1977; Middlebrooks et al. 1980). There is additional – although in some cases only preliminary – evidence that other response properties are also spatially segregated or systematically organized within AI. Among those properties are aspects of responses related to sound localization (e.g., Imig et al. 1990; Jenkins and Merzenich 1984; Middlebrooks and Pettigrew 1981; Rajan et al. 1990; Reale and Kettner 1986), inhibitory interaction (Shamma and Fleshman 1990; Suga and Tsuzuki 1985), sharpness of tuning (Schreiner and Cynader 1984; Schreiner and Mendelson 1990; Sutter and Schreiner 1991a), and the growth of response strength (Phillips et al. 1985; Reale et al. 1979).

In an attempt to assess the global distribution and possible covariances of functional characteristics within AI, 12 properties of multiple-unit responses along the dorsoventral extent of cat AI were investigated (Schreiner et al. 1988; Mendelson et al. 1988). In an initial report of this experimental series, the spatial distribution of the "integrated excitation" as reflected in the sharpness of tuning curves and the response strength to

Correspondence to: C. Schreiner

broadband transients was described (Schreiner and Mendelson 1990). A consistent systematic distribution of the spectrally integrated excitation along the isofrequency domain was observed, providing further evidence for a systematic internal organization of AI. (Note that the term "isofrequency" is used here to describe a hypothetical line along the cortical dimension that is oriented orthogonal to the main CF gradient. Whether indeed there are true isofrequency contours across AI containing neurons of exactly the same CF has never been shown.)

The current report describes the spatial distribution of neural response properties relating to the representation of pure tones as a function of intensity for the same cortical locations that were studied regarding excitatory bandwidth and responses to transients (Schreiner and Mendelson 1990). Several aspects of the changes in firing rate with stimulus level as they are reflected in rate/level functions of multiple-unit responses were utilized in further exploring the spatial organization of AI in the anesthetized cat. The studied parameters are response threshold, dynamic range, monotonicity, and strongest response level (SRL). In particular, the relationship of the spatial distribution of intensity parameters with that of the excitatory bandwidth and responses to broadband stimuli (Schreiner and Mendelson 1990) will be discussed.

Materials and methods

Surgical preparation

The methods are the same as described in previous, related reports (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a). Briefly, results were obtained in the right hemispheres of adult cats. Anesthesia was induced with an intramuscular injection of ketamine hydrochloride (10 mg/kg) and acetylpromazine maleate (0.10 mg/kg). Animals were maintained at a surgical level of anesthesia with a continuous infusion of pentobarbital sodium (2 mg/kg per hour) in lactated Ringer's solution. The temperature of the animals was monitored with a rectal probe and maintained at 37.5° C by means of a heated water blanket with feedback control.

The head was fixed and the temporal muscle on the right hemisphere was then retracted and the lateral cortex exposed by a craniotomy. The dura overlying the middle ectosylvian gyrus was removed, the exposed cortex covered with silicone oil, and a photograph of the surface vasculature taken to record the electrode penetration sites.

Stimulus generation and delivery

Experiments were conducted in a sound-shielded room (Industrial Acoustics Company). Auditory stimuli were presented via calibrated headphones (STAX 54) enclosed in small chambers that were connected to sound delivery tubes sealed into the acoustic meatuses (Sokolich, US Patent 4251686; 1981). The frequency response of the system was essentially flat up to 12 kHz and did not have major resonances deviating more than 6 dB from the average level. Above 15 kHz, the output rolled off at a rate of 10 dB/octave.

Tones were generated by a microprocessor (TMS32010; 16-bit digital-analog (D/A) converter at 120 kHz; low-pass filter at 35 kHz or 50 kHz). Additional attenuation was provided by a pair of passive attenuators (Hewlett-Packard). The duration of the tone bursts was 50 ms including 3 ms rise/fall time. The interstimulus interval was 500–1000 ms.

For each recording site, responses were recorded to at least 6–5 different tone bursts. Tone bursts were presented in a pseudorandom sequence of different frequency/level combinations selected from 15 level values (70 dB range) and 45 frequency values (equidistant on a logarithmic frequency scale; 2–4 octaves range). From the responses to all stimuli, frequency response areas (FRAs) were reconstructed.

Recording procedure

Parylene-insulated tungsten microelectrodes with impedances at 1 kHz of 0.8–1.3 MΩ were used. All penetrations were roughly orthogonal to the brain surface. The recordings were derived at an intracortical depth ranging from 600 to 1000 μm, as determined by the microdrive setting, roughly corresponding to cortical layers III and IV. Neuronal activity of single units or small groups of neurons (2–6 units) were amplified, band-pass filtered, and monitored on an oscilloscope and an audio monitor. Spike activity was separated from the background noise with a level discriminator (BAK DIS-1) set at more than 50% of the background noise and stored in a computer (DEC 11/73). The recording window had a duration of 50 ms, corresponding to the stimulus duration.

Data analysis

From the FRAs, response strengths to the CF and the two stimulus frequencies closest to the CF were extracted. Spike count level profiles were constructed for each recording site by summing the spike counts for each signal level produced by these frequencies, i.e., responses from 45 signal presentations (15 levels for three frequencies) were utilized. The following response parameters were obtained:

A. *Contralateral threshold*, the lowest level of a contralateral tone (expressed in dB SPL) that reliably evoked a response. This response parameter was directly determined from audiovisual cues of the raw response during stimulation with contralateral tones while manually controlling stimulus frequency and level. This measure was used for the representation of the response threshold.

From the rate/level functions, the following response parameters were extracted:

B. *Transition point*, the point in the rate/level function that marked the transition from a fast-growing, low-level portion to a less fast growing, saturating or decreasing high-level portion of the rate level function. In virtually all cases, a monotonic, fast-growing portion could be distinguished that rose from the response threshold (see Fig. 1, open arrows) to a transition point (see Fig. 1, filled arrow), i.e., a clear transition to a less fast growing portion of the rate level profile, a saturation, or to a decline of firing rate.

C. *Dynamic range*, the monotonic, fast-growing, low-level portion of a rate/level function. The dynamic range was rounded to the nearest multiple of 5 dB.

D. *SRL*, the stimulus level (in dB SPL) evoking the highest firing rate in the range covered by the rate/level profile. If the rate level function was monotonically growing up to the highest level used for the reconstruction, the highest level was used as an estimate.

E. *Monotonicity*, the rate of change in the rate/level profile for stimulus levels above the transition point of the function. To quantitatively determine the degree of rate change or the amount of monotonicity above the transition point, the high-level slope of the rate/level profile was estimated by a linear regression analysis. The starting point of the regression analysis was at the transition point of the rate/level profile (Fig. 1, filled arrow). The end point was either the rate at the highest stimulus level used to construct the profile or the stimulus level at which the firing rate fell below 90% of the firing rate of the transition point minus spontaneous rate. A minimum of five level values, including the transition point itself,

#706/28   DR=20 dB
CF=11.4 kHz SRL=40 dB
Thr.=20 dB Mon=-2.14

#706/70   DR=10 dB
CF=14.3 kHz SRL=10 dB
Thr.=0 dB Mon=-1.14

#706/42   DR=15 dB
CF=10.5 kHz SRL=20 dB
Thr.=5 dB Mon=0.0

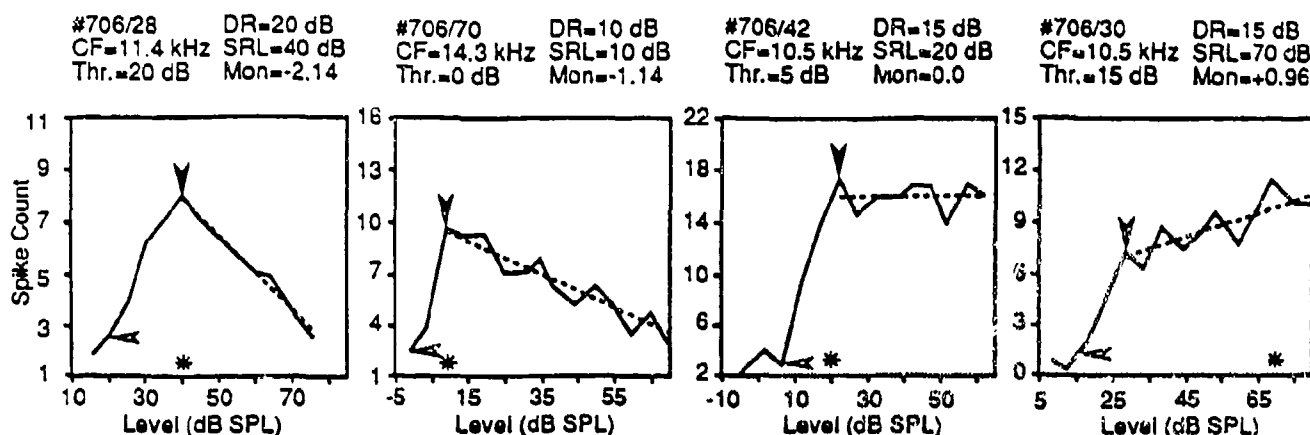#706/30   DR=15 dB
CF=10.5 kHz SRL=70 dB
Thr.=15 dB Mon=+0.96

Fig. 1. Spike count level functions of primary auditory cortex (AI). Four representative examples of spike count level functions are shown. All were obtained in the same animal (87–706). Response thresholds are indicated by open arrows. The transition point (see Materials and methods) is marked by a *filled arrow*. The strongest response level (*SRL*) of each location, i.e., the stimulus level that produced the highest firing rate, is indicated by a *star*. The *dashed line* in the high-level portion of the function is the linear regression estimate of the spike count as a function of level. The slope or monotonicity (*Mon*) of the function is indicated above each example expressed in percentage per decibels (% dB: 100% corresponds to the firing rate at the transition point). Characteristic frequency (*CF*), threshold (*Thr.*; in dB SPL), dynamic range (*DR*, level range between threshold and transition point), and SRL (dB sound pressure level SPL) are also given for each example

had to be included in the analysis, or the cortical location was excluded from further consideration of this parameter. The slope of the resulting regression function (percentage decibels: Fig. 1, dashed lines) was taken as measure of the monotonicity of each recording site.

Data representation

Pseudo three-dimensional projections and contour plots were utilized to represent the spatial distribution of response parameters across the cortical surface (topography software from Golden Software; see Schreiner and Mendelson 1990). The actual spatial locations of the recording sites were used to generate a two-dimensional grid of the represented area by projecting the actual sites to the nearest grid point. Elevation of the grid surface corresponded to the spatially averaged magnitude of a functional parameter at a given site. The grid program employed a weighted inverse distance squared algorithm for the interpolation of all grid points. The result of the transformations provided a faithful although slightly amplitude compressed picture of the spatial distribution of the original information (see Fig. 3 in Schreiner and Mendelson 1990).

Since the pseudo three-dimensional projection of the maps always resulted in some perspective distortions, an undistorted view of the spatial distribution is provided by presenting detailed contour plots (line segments that connect interpolated locations with equal surface elevation or z-values) of one of the three cases. To judge the correspondence of the data values with their three-dimensional representation, the value and location of each data point is provided for one exemplary case. The distribution of recording locations for the other two maps as well as the corresponding distributions of characteristic frequency, sharpness of tuning, and transient responses can be obtained from the maps shown in the first part of this study (Schreiner and Mendelson 1990).

Results

Response threshold

Response thresholds to contralateral stimulation with CF tones were obtained for 261 locations in the high-

frequency domain of AI for CFs between 5.8 and 26.3 kHz. AI was identified by its rostrocaudal CF gradient. The dorsoventral extent of the mapping was carried out until clear deviations from the cochleotopic organization were encountered, which were presumed to indicate that the end of AI was reached (Middlebrooks and Zook 1983; Schreiner and Cynader 1984) The locations of the mapped areas on the right hemisphere of cat cortex are schematically shown in Fig. 2 (lightly shaded regions). The orientation of the isofrequency domain (thin line) is indicated by an estimated isofrequency contour (for details see Schreiner and Mendelson 1990, Fig. 4). The functional border to the second auditory field (AII; Fig. 2, dashed line) was estimated from the local distribution of CFs, the sharpness of tuning, and the response threshold (Schreiner and Cynader 1984)

The distributions of the encountered response thresholds in the studied cortices is shown as a function of CF in Fig. 3. For any given CF, a range of response thresholds was observed that extended over 20 dB to more than 50 dB. Mean and standard deviation of the threshold values are given in Table 1 for each case separately. For each animal, the lowest response thresholds were between −10 and 10 dB SPL, indicating a similar degree of sensitivity.

Reconstructions of the spatial distribution of response thresholds clearly indicated that the obtained range of thresholds was not randomly distributed across the mapped area. Figure 4A illustrates the three-dimensional reconstruction of case 87–518, showing high response thresholds in the ventral portion of the mapped area, systematically decreasing thresholds toward a region with predominantly low thresholds in the middle of the map, and a moderate increase in response threshold toward the dorsal third of the map. This latter area of the map showed some regional variability in the threshold values. The relationship between the orientation of iso-
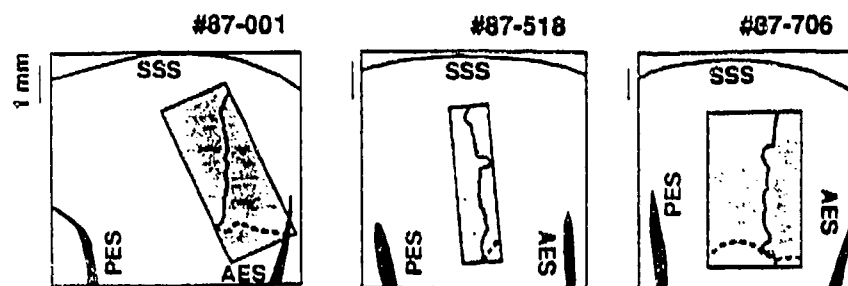
#87-001   #87-518   #87-706

Fig. 2. Schematic representation of the mapped areas on the right hemisphere of cat cortex. The *lightly shaded rectangle* approximates the mapped area (see Fig. 4 in Schreiner and Mendelson 1990 for more details). The orientation of the isofrequency axis is indicated by the course of one representative isofrequency contour *(thin line)*. The *contours* correspond to 18 kHz (#87-001), 10 kHz (#87-518), and 11 kHz (#87-706), respectively. The *dashed lines* indicate the estimated border between primary auditory cortex (AI) – located dorsal to the border – and secondary auditory cortex (AII). *SSS*, suprasylvian sulcus; *AES*, anterior ectosylvian sulcus; *PES*, posterior ectosylvian sulcus
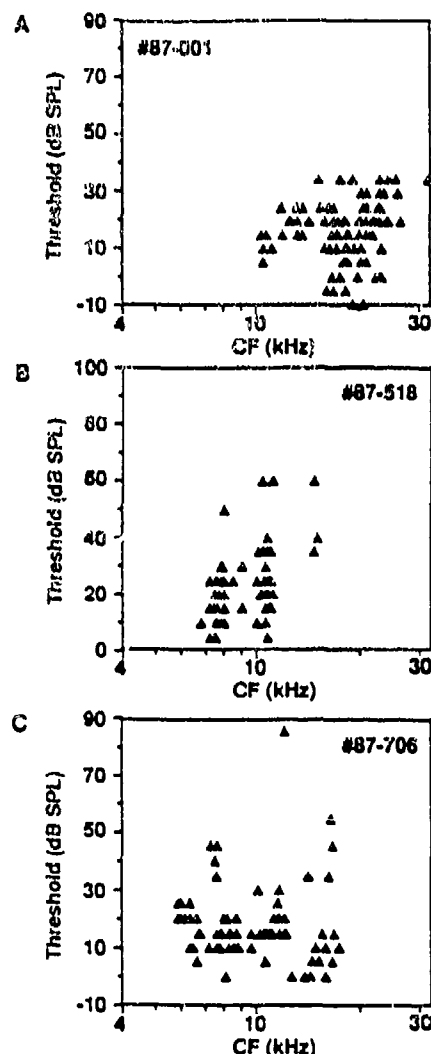

Fig. 3A–C. Contralateral response thresholds. All encountered threshold values are plotted as a function of characteristic frequency (*CF*) for each of the three cases. Mean thresholds with standard deviations are given in Table 1

frequency contours (see thin dashed line in the contour plot of Fig. 4A) and the overall pattern in the threshold distribution might suggest that response thresholds vary systematically along the isofrequency domain. In this case, at least the middle and ventral portion of the map suggest that the global threshold distribution pattern may be independent of the cochleotopic gradient. The area of lowest thresholds includes the location of the reversal in the dorsoventral gradient of the integrated excitatory bandwidth (thick dashed line; mean from Q-10 dB and Q-40 dB; Schreiner and Mendelson 1990) which, correspondingly, represents the most sharply tuned area along the isofrequency axis of AI.

Aggregates of locations with higher or lower response thresholds were also seen in the two other cases as demonstrated in Fig. 4B, C. Both maps, however, showed a more complex distribution of high- and low-threshold aggregates. In particular, the spatial frequency (spatial rate of change between maxima and minima) of the threshold distribution of the two cases appeared to be higher than in case 87-518. High- and low-threshold aggregates were scattered throughout the entire extent of the mapped areas and showed no clear spatial alignment with the tonotopic gradient or the isofrequency axis. The shape of areas with similar thresholds varied widely. Finally, the functional/spatial transitions from, for example, a low-threshold to a high-threshold aggregate varied and could be either gradual or quite steep.

In summary, a large range of threshold values was found even for responses with similar CF. Spatial aggregates could be delineated in AI that covered the same, relatively narrow ranges of response thresholds. The spatial extent and location of aggregates with similar response thresholds appeared to vary greatly from animal to animal, and varied with regard to their locations in the isofrequency domain of AI.
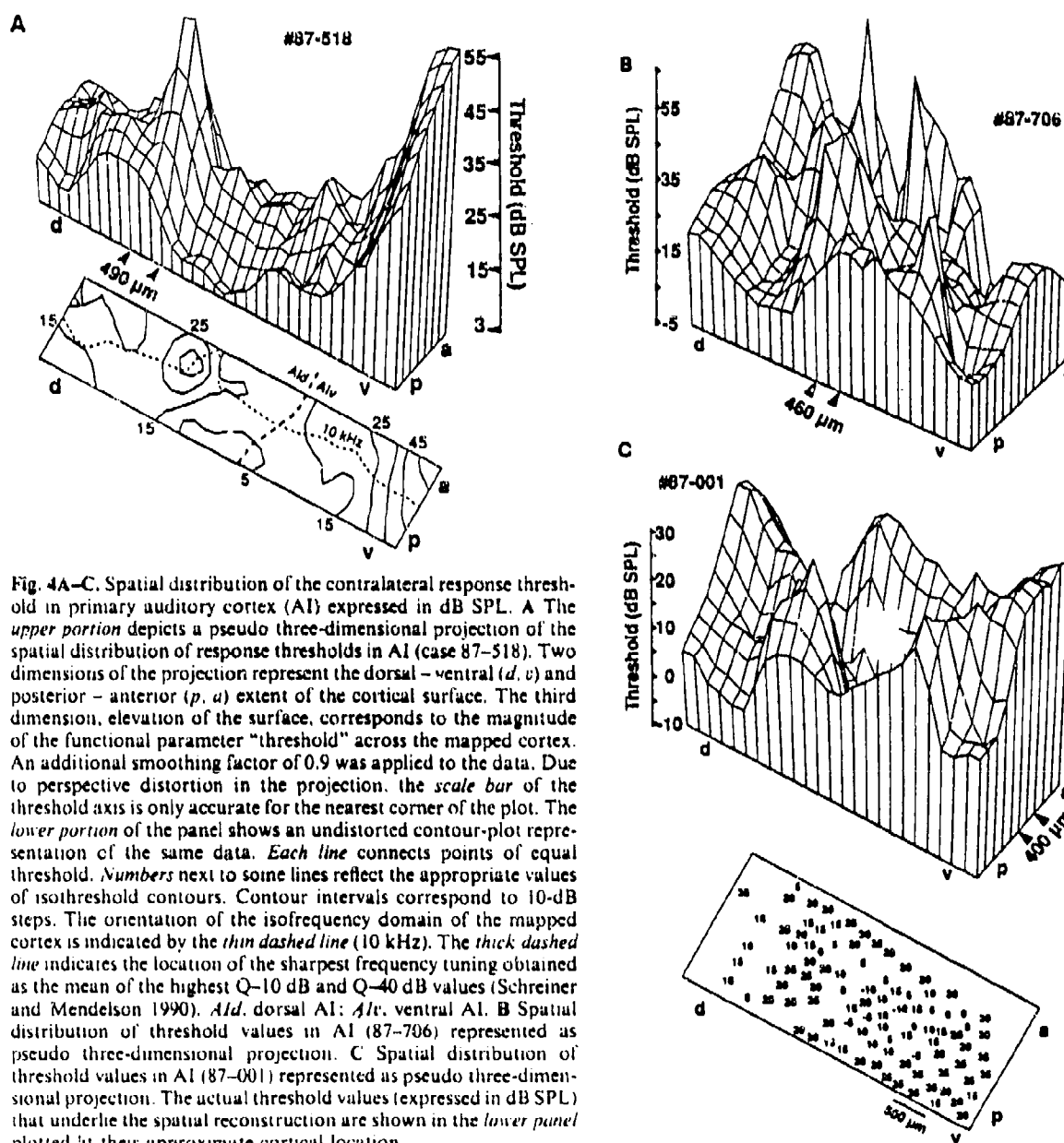
Dynamic range

Examples of the dynamic range of rate/level functions, as defined above (see Materials and methods), are presented in Fig. 1 as the intensity range between the

Table 1. Statistical description of response threshold. dynamic range. strongest response level. and monotonicity for cortical rate level functions

	Case					
	87–001		87–518		87–706	
	Mean ± SD	Cortical locations (n)	Mean ± SD	Cortical locations (n)	Mean ± SD	Cortical locations (n)
Threshold (dB SPL)	16.3 ± 10.6	95	22.2 ± 12.2	81	16.9 ± 10.9	85
Dynamic range (dB)	20.8 ± 6.5	86	19.9 ± 8.4	68	16.6 ± 5.9	80
SRL (dB SPL)	35.8 ± 11.0	86	49.4 ± 17.1	77	39.4 ± 16.9	82
Monotonicity (% dB)	−0.14 ± 0.76	86	−0.60 ± 0.90	67	−0.30 ± 0.57	80

SRL. strongest response level

Fig. 4A–C. Spatial distribution of the contralateral response threshold in primary auditory cortex (AI) expressed in dB SPL. A The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of response thresholds in AI (case 87–518). Two dimensions of the projection represent the dorsal – ventral (*d, v*) and posterior – anterior (*p, a*) extent of the cortical surface. The third dimension, elevation of the surface, corresponds to the magnitude of the functional parameter "threshold" across the mapped cortex. An additional smoothing factor of 0.9 was applied to the data. Due to perspective distortion in the projection. the *scale bar* of the threshold axis is only accurate for the nearest corner of the plot. The *lower portion* of the panel shows an undistorted contour-plot representation of the same data. *Each line* connects points of equal threshold. *Numbers* next to some lines reflect the appropriate values of isothreshold contours. Contour intervals correspond to 10-dB steps. The orientation of the isofrequency domain of the mapped cortex is indicated by the *thin dashed line* (10 kHz). The *thick dashed line* indicates the location of the sharpest frequency tuning obtained as the mean of the highest Q–10 dB and Q–40 dB values (Schreiner and Mendelson 1990). *AId.* dorsal AI; *AIv.* ventral AI. B Spatial distribution of threshold values in AI (87–706) represented as pseudo three-dimensional projection. C Spatial distribution of threshold values in AI (87–001) represented as pseudo three-dimensional projection. The actual threshold values (expressed in dB SPL) that underlie the spatial reconstruction are shown in the *lower panel* plotted at their approximate cortical location
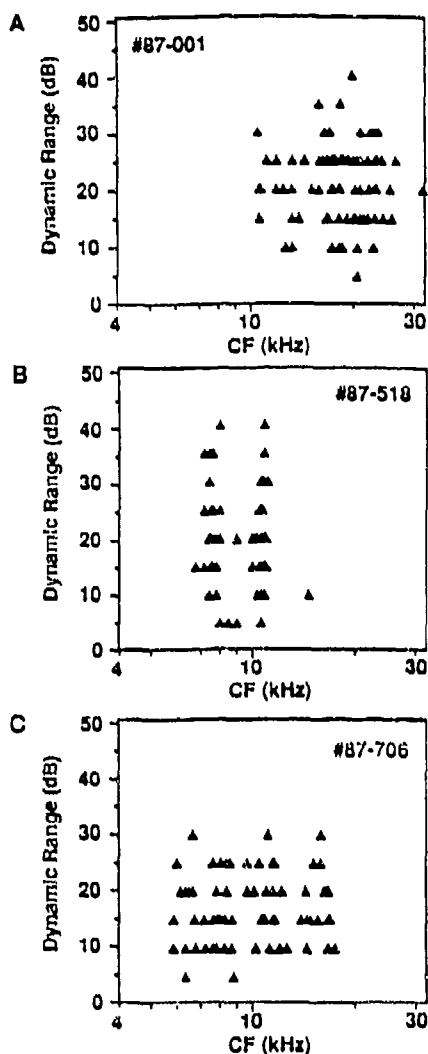
A

B

C

Fig. 5A–C. Dynamic range of rate-level functions. All encountered dynamic range values are plotted as a function of characteristic frequency (CF) for each of the three cases. Mean values with standard deviations are given in Table 1

open and closed arrows. At all recording sites for which a dynamic range could be defined, values were between 5 and 40 dB. Mean and standard deviations of the encountered dynamic ranges are given in Table 1. The spread of values was essentially independent of the CF (see Fig. 5).

The spatial distribution of the dynamic range values, as shown in Fig. 6, appeared to be quite idiosyncratic. Spatial aggregates of several recording locations with similar dynamic ranges were identified. However, the representation of similar dynamic ranges was quite fragmented, leading to steep gradients between aggregates with large or small values. Consequently, a relatively high spatial frequency in the distribution of the functional parameter range was observed. This attribute of rate level functions did not show a clear systematic spatial distribution that appeared to be aligned with either the CF gradient or the isofrequency axis and differed substantially from case to case.

Monotonicity of rate-level functions

The distribution of monotonicity values. i.e.. the approximated slope of rate-level functions above the transition point (see Materials and methods), are shown in Fig. 7. A quantitative evaluation of the monotonicity or nonmonotonicity of the obtained rate level functions revealed that 60.5% of the 233 locations for which the analysis could be completed were nonmonotonic (slopes < 0% dB). Strongly nonmonotonic rate level functions, arbitrarily defined as having slopes of less than −1%·dB, were obtained for 19.3% of all recorded locations or 31.9% of all rate level functions with negative slopes.

The spatial distribution of the degree of monotonicity is shown in Fig. 8. The elevation of the grid surfaces corresponds to negative slope values, and therefore expresses the degree of nonmonotonicity of rate level functions: increasing elevation represents increased values of nonmonotonicity. In cases 87–518 (Fig. 8A) and 87–706 (Fig. 8B), two prominent, circumscribed regions with medium to high degrees of nonmonotonicity can be delineated. The ventral area of nonmonotonicity appeared in both cases to be spatially more extended than the dorsal region. Both regions were separated by a relatively narrow region containing locations with positive or near-zero slope values. The degree of monotonicity varied substantially and systematically in the isofrequency domain, whereas the variations in the spatial dimension of the CF gradient, i.e., orthogonal to the isofrequency domain, appeared to be smaller and lacked a clear systematic organization.

The third case, 87–001, showed a more complex spatial distribution of the monotonicity of rate level functions that appeared to differ somewhat from the other two cases. However, the isofrequency contour was not aligned with the dorsoventral extent of the mapped area but approximated the posterior–ventral·anterior–dorsal diagonal of the map (Schreiner and Mendelson 1990; see also Figs. 2, 11). Consequently, similar interpretations of the spatial distribution can be derived as outlined for the other two cases, i.e., an orderly progression in the isofrequency domain and orthogonality to the frequency gradient. The only remaining difference between this and the other two cases is that no clearly expressed dorsal aggregate of nonmonotonicity was observed, although several smaller patches in the dorsal part of AI appear to be highly nonmonotonic. All three cases showed that the ventral region of high nonmonotonicity was closely aligned with the region of sharpest frequency tuning (see thick dashed lines in Figs. 8A and 11).

Strongest response level

The stimulus level producing the highest firing rate for a CF tone (SRL), varied between 10 and 80 dB SPL, as shown in Fig. 9. For each CF, essentially the full range of SRLs was encountered, showing a fairly even distribution across the level range. Mean and standard deviations are given in Table 1.
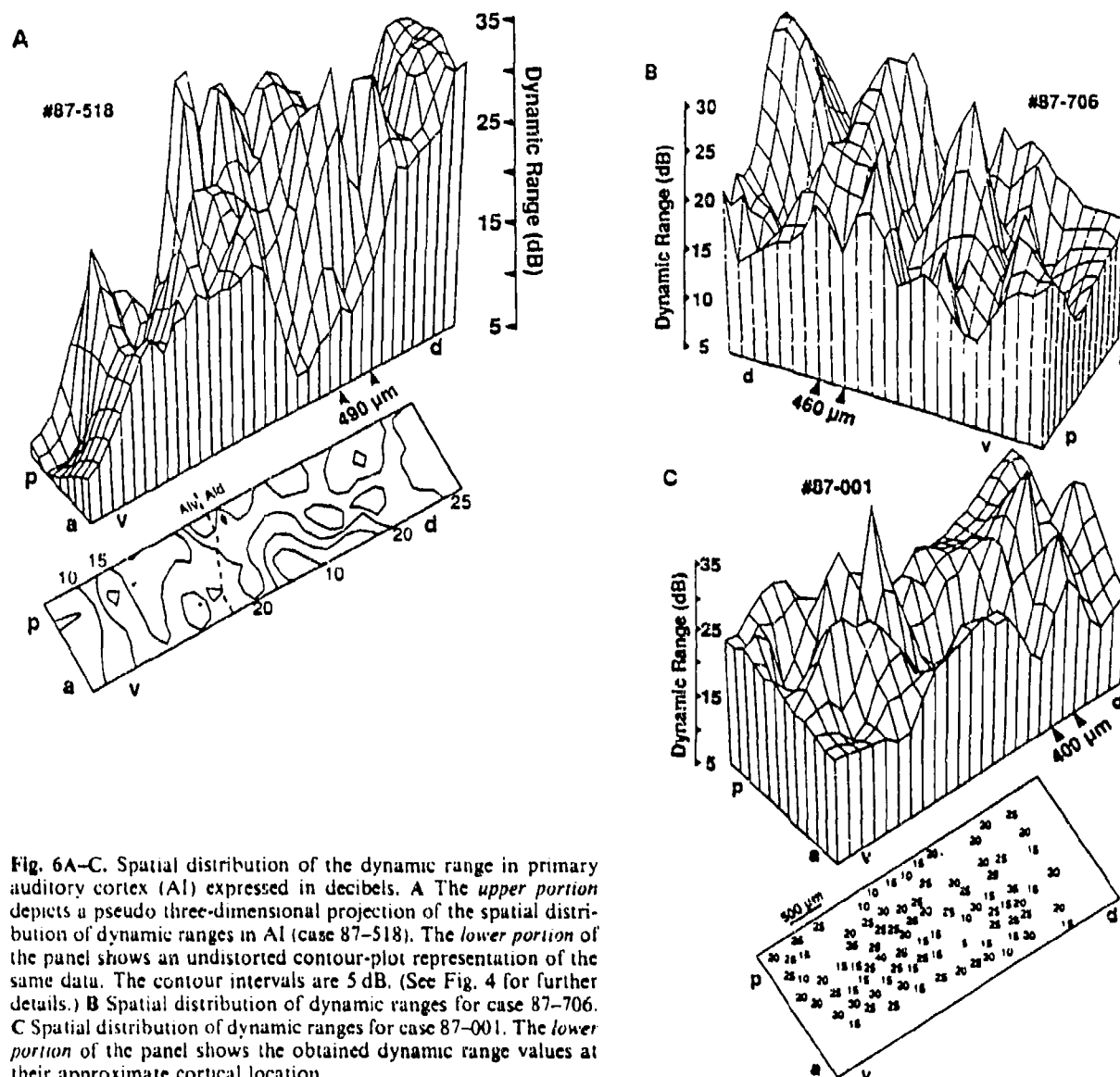
Fig. 6A–C. Spatial distribution of the dynamic range in primary auditory cortex (AI) expressed in decibels. A The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of dynamic ranges in AI (case 87–518). The *lower portion* of the panel shows an undistorted contour-plot representation of the same data. The contour intervals are 5 dB. (See Fig. 4 for further details.) B Spatial distribution of dynamic ranges for case 87–706. C Spatial distribution of dynamic ranges for case 87–001. The *lower portion* of the panel shows the obtained dynamic range values at their approximate cortical location

The spatial distribution of the SRL values is shown in Fig. 10. In all three cases, spatial aggregates of recording locations with the lowest SRLs were encountered in the dorsoventral center of the mapped areas. In case 87–518, the SRLs systematically increased toward the ventral and dorsal boundaries of AI. For the other two cases, spatial aggregates of locations with medium and high SRLs were evident; however, each spatial distribution differed substantially from the other cases and showed no clear pattern that appeared to be aligned with either the CF gradient or the isofrequency axis. Note that some of the lowest SRLs were often found near the region with sharpest frequency tuning (see thick dashed line in Figs. 10A, 11).

Covariance of rate/level function and spectral characteristics

The parameters that were extracted from the rate level function to describe the influence of stimulus level on the response characteristic of cortical locations were not necessarily independent of each other. The SRL of a given cortical location, for example, is in essence proportional to the response threshold, the dynamic range, and the monotonicity of the rate/level function. A correlational analysis of the studied parameters reflects these interdependencies. Tables 2–5 show all parameter combinations that were correlated at a level of significance of $P < 0.05$. In addition, the correlations of these intensity parameters with those of the spectral parameters integrated excitatory bandwidth and transient response are given.
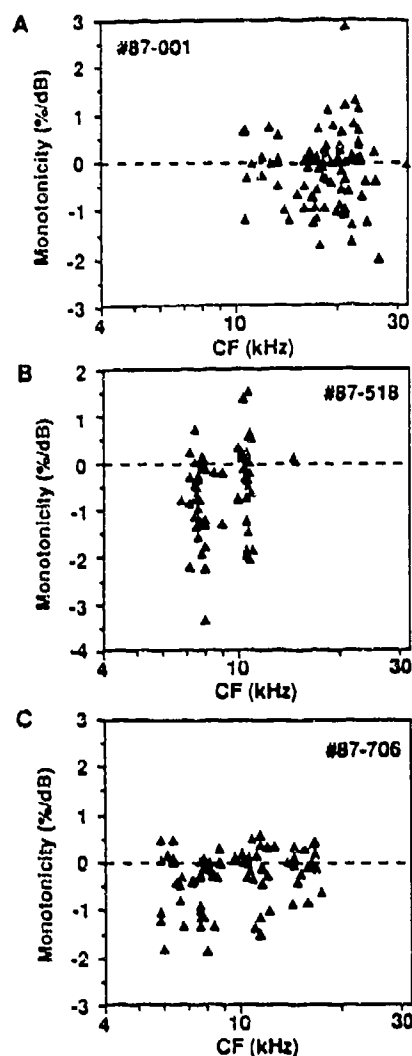
Fig. 7A–C. The monotonicity measure of the high-level portion of rate level functions (see Materials and methods). All encountered monotonicity values are plotted as a function of characteristic frequency (*CF*) for each of the three cases. The *dashed line* separates locations with monotonic rate/level functions from locations with nonmonotonic rate level functions. Mean values with standard deviations are given in Table 1

Response thresholds (Table 2) showed the highest and, across cases, most reliable correlation with the SRL. In two of the three cases this correlation was most strongly expressed in the ventral aspects of AI. A weak but fairly consistent negative correlation was evident with the excitatory bandwidth, especially with Q–40 dB in AIv, i.e., the higher the threshold, the broader was the frequency tuning. Particularly in case 87–001, the strength of the step response was correlated with the response threshold.

Monotonicity values were consistently correlated with the SRLs across the entire extent of AI (Table 3). Equally reliable correlations were seen between the monotonicity and Q–40 dB across AI, but especially in AIv. In addition, monotonicity was positively correlated with the strength of step responses across all of AI, with the exception of the dorsal aspects of case 87–706.

In addition to correlations with response threshold and monotonicity, the SRLs showed strongest correlations with the step responses, particularly in AIv (Table 4). Correlation between other parameters were either weak or not consistent across cases.

Spatial alignment of intensity and spectral parameters

To compare the spatial distribution of the studied response characteristics with the location of the area of sharpest tuning in the same cases (Schreiner and Mendelson 1990), contour plots of the three cases are given in Fig. 11. Shaded areas in these contour plots represent approximately the spatial locations of the lower half of the encountered parameter range for response threshold, dynamic range, and SRL (see Table 1). For the representation of monotonicity values a different criterion was chosen, namely that of moderate to high nonmonotonicity. Accordingly, shaded areas correspond to values below $-0.5\%/dB$ for cases 87–001 and 87–706, and below -0.4% dB for case 87–518.

Comparison of the spatial distributions of level-related response parameters within and across animals suggests a high degree of idiosyncrasy in these aspects of functional organization of auditory cortex. The spatial pattern of monotonicity appeared to be the most consistently organized of the four parameters described here. A circumscribed area of moderate to high nonmonotonicity was reliably encountered in the center of high-frequency AI along the main CF gradient and coincided with the area of narrowest integrated excitatory bandwidth. These elongated patches of high nonmonotonicity were arranged orthogonal to the isofrequency domain and, although they showed interruptions, appeared to cover the entire studied frequency range. Smaller, patchy aggregates of nonmonotonic rate level functions were consistently encountered dorsal to the main area of nonmonotonicity. The spatial distribution of the SRL measure showed a similar, although slightly more complex, pattern. Since monotonicity and SRL were the most strongly correlated parameters, this spatial covariance was to be expected.

For all three cases, the response threshold values were moderately correlated with SRL values (see Table 2). The spatial distribution of response thresholds appeared to be more erratic and showed some deviation from the general pattern of alignment with the CF gradient and the sharpness of tuning gradient as seen for monotonicity and, to a lesser degree, SRL.

Finally, the spatial pattern derived for the dynamic range of rate/level functions showed the least resemblance to the spatial distributions of the other parameters under consideration, thus reflecting the virtual absence of correlation with those parameters (Table 5).
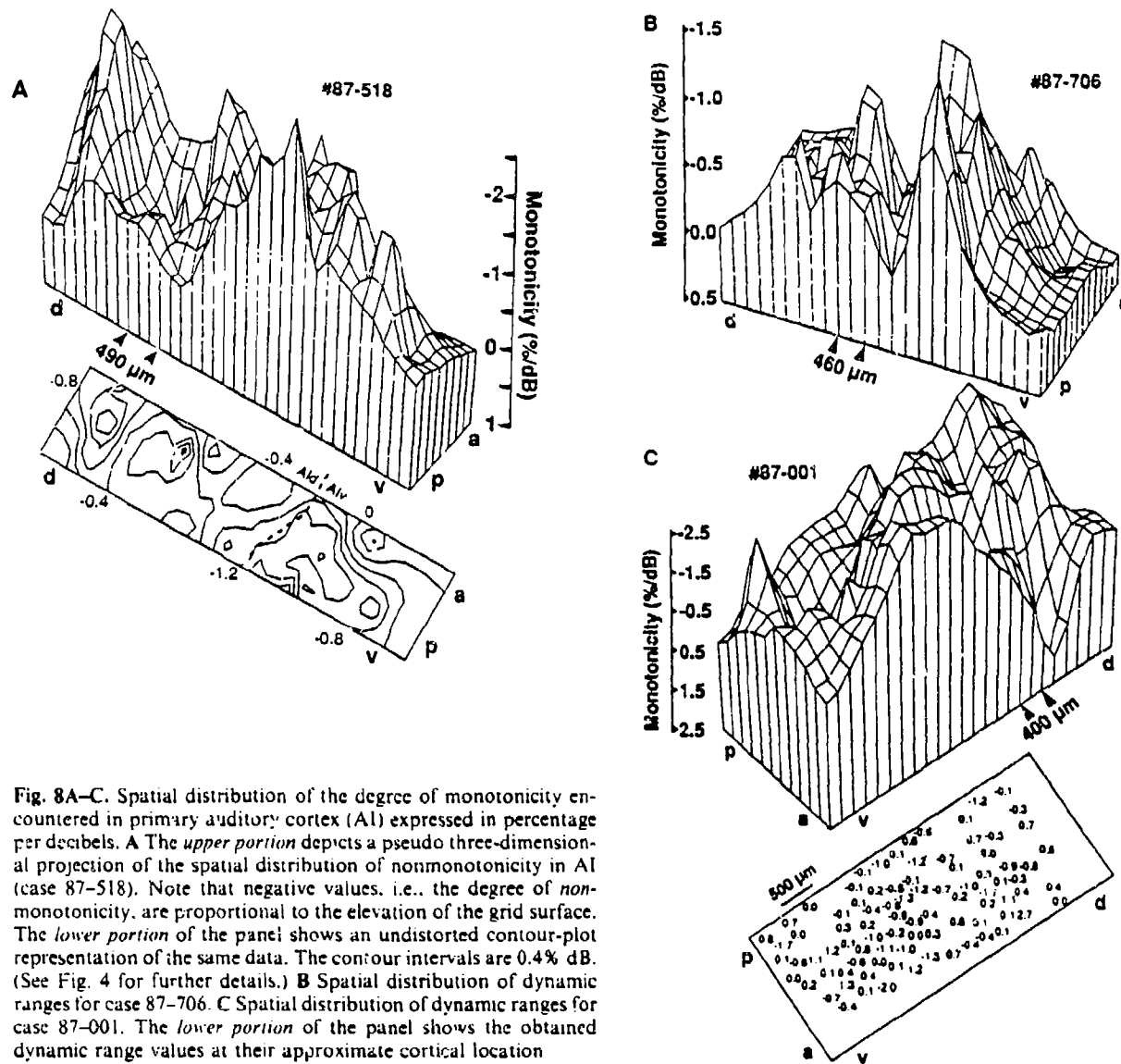
Fig. 8A–C. Spatial distribution of the degree of monotonicity encountered in primary auditory cortex (AI) expressed in percentage per decibels. **A** The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of nonmonotonicity in AI (case 87–518). Note that negative values, i.e., the degree of *non*-monotonicity, are proportional to the elevation of the grid surface. The *lower portion* of the panel shows an undistorted contour-plot representation of the same data. The contour intervals are 0.4% dB. (See Fig. 4 for further details.) **B** Spatial distribution of dynamic ranges for case 87–706. **C** Spatial distribution of dynamic ranges for case 87–001. The *lower portion* of the panel shows the obtained dynamic range values at their approximate cortical location

Table 2. Correlation of contralateral minimum threshold with cortical response measures obtained at the same locations

	Case	AI		AId[a]		AIv[b]	
		r	P	r	P	r	P
MON	87–518	0.32	0.01			0.56	0.0001
SRL	87–516	0.62	0.0001			0.78	0.0001
	87–706	0.35	0.001			0.33	0.04
	87–001	0.48	0.0001	0.55	0.0001		
DR	87–001					−0.49	0.006
Q–10 dB	87–001	−0.29	0.004	−0.28	0.03		
Q–40 dB	87–518	−0.26	0.02			−0.34	0.05
	87–001	−0.25	0.02			−0.48	0.003
Step	87–516			0.35	0.04		
	87–001	0.36	0.0004	0.26	0.05	0.41	0.02

Linear regression analysis; r = correlation coefficient; P = level of significance (F-test)
[a] AId represents recording location in the portion of AI dorsal to the most sharply tuned multiple unit responses (Schreiner and Mendelson 1990)
[b] AIv represents location ventral to the most sharply tuned region of AI
MON, monotonicity; SRL, strongest response level; DR, Dynamic range; Step, response strength to a rapid frequency step from 0.2 to 64 kHz, i.e., broadband transient (Schreiner and Mendelson 1990)
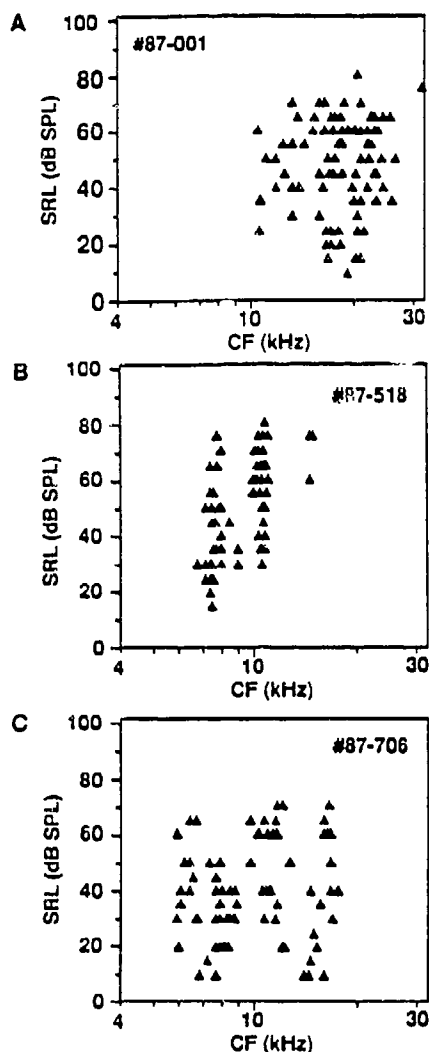
Fig. 9A–C. The strongest response level (*SRL*) of rate level functions (see Materials and methods). All encountered SRL values are plotted as a function of characteristic frequency (*CF*) for each of the three cases. Mean values with standard deviations are given in Table 1

Discussion

Response thresholds and rate level functions for CF tones were determined in the primary auditory cortex of anesthetized cats using the multiple-unit recording technique and related to the estimates of integrated excitatory bandwidth obtained at the same locations. Three descriptive and functionally significant parameters were extracted from the rate-level functions, namely dynamic range, SRL, and monotonicity. The spatial distribution of these parameters across the high-frequency portion of AI was then determined. In addition to the four parameters described in this report, eight other parameters were studied at the same locations in the same animals in order to determine the global distribution of response properties in AI. A major goal of this study was to compare and, potentially, relate the distributions of the obtained response properties to each other. Since it is difficult to obtain maps for such a large number of stimulus/response parameters, only the three most complete cases were included in this and a previous report that described the spatial distribution of four spectral parameters (CF, Q–10 dB, Q–40 dB, and transient response strength; Schreiner and Mendelson 1990). Interpretations and discussions of methodological influences will be given for each of the studied parameters separately followed by a brief comparison with the spatial distributions of other parameters.

Response threshold

The obtained range of contralateral response thresholds of approximately 50 dB encountered in this study is in close agreement with threshold ranges seen in previous multiple-unit (Schreiner and Cynader 1984) and single-unit studies (Phillips and Irvine 1981) of AI. In two of the three cases, the lowest thresholds were essentially independent of CF in the tested ranges of 0.65 and 1.57 octaves. In the third case (87–001), CFs between 12 and 14 kHz had thresholds approximately 25 dB above the

Table 3. Correlation of monotonicity of rate level functions with cortical response measures obtained at the same locations

	Case	AI		AId		AIv	
		r	P	r	P	r	P
THR	87–518	0.32	0.01			0.56	0.0001
SRL	87–518	0.46	0.0001	0.36	0.03	0.41	0.04
	87–706	0.58	0.0001	0.63	0.0001	0.51	0.002
	87–001	0.54	0.0001	0.51	0.0001	0.52	0.003
DR	87–706	−0.24	0.03			−0.47	0.004
Q–10 dB	87–001	−0.29	0.004	−0.28	0.03		
Q–40 dB	87–518	−0.50	0.001	−0.35	0.03	−0.56	0.003
	87–706	−0.29	0.008			−0.34	0.05
	87–001	−0.25	0.02			−0.48	0.003
Step	87–518	0.50	0.0001	0.53	0.0006	0.43	0.03
	87–706					0.49	0.004
	87–001	0.39	0.0002	0.35	0.01	0.48	0.008

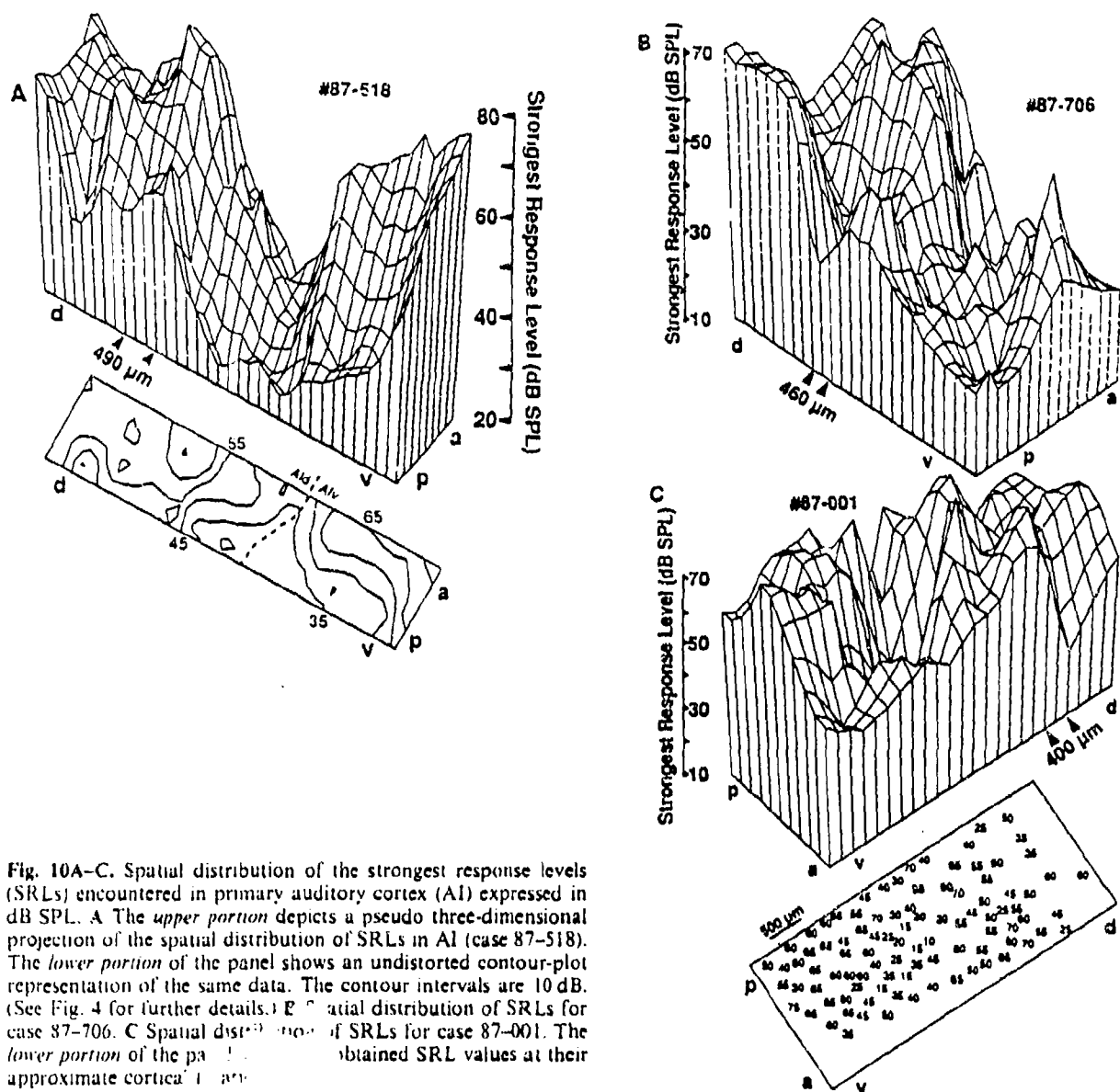THR, minimum response threshold. For further explanations see Table 2

Fig. 10A–C. Spatial distribution of the strongest response levels (SRLs) encountered in primary auditory cortex (AI) expressed in dB SPL. A The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of SRLs in AI (case 87-518). The *lower portion* of the panel shows an undistorted contour-plot representation of the same data. The contour intervals are 10 dB. (See Fig. 4 for further details.) B atial distribution of SRLs for case 87-706. C Spatial dist   of SRLs for case 87-001. The *lower portion* of the pa   btained SRL values at their approximate cortical

Table 4. Correlation o strongest response level with cortical response measures obtained at the same locations

|	Case	AI		AId		AIv	
		r	P	r	P	r	P
THR	87-516	0.o2	0.0001			0.78	0.0001
	87-706	0.35	0.001			0.33	0.04
	87-001	0.48	0.0001	0.55	0.0001		
MON	87-518	0.46	0.0001	0.36	0.03	0.41	0.04
	87-706	0.58	0.0001	0.63	0.0001	0.51	0.002
	87-001	0.54	0.0001	0.51	0.0001	0.52	0.003
Q–10 dB	87-706	−0.28	0.01			−0.47	0.003
Q–40 dB	87-518	−0.28	0.02			−0.42	0.02
	87-706	−0.35	0.001				
Step	87-518	0.26	0.02			0.45	0.008
	87-706	0.26	0 02			0.60	0.0001
	87-001	0.42	0.0001	0.39	0.003	0.47	0.005

For explanations see Tables 2, 3
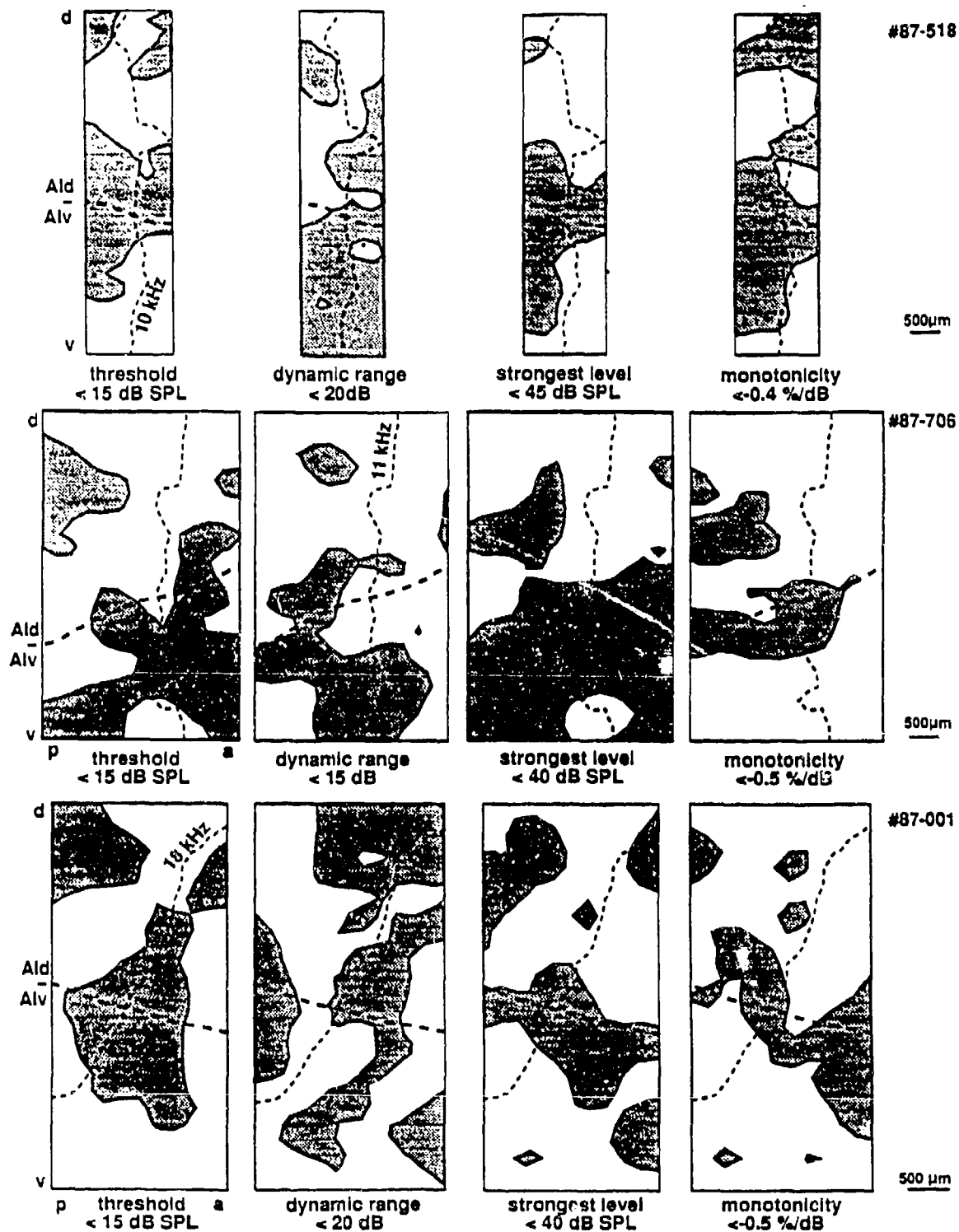
Fig. 11. Comparison of the spatial distribution of response parameters in primary auditory cortex (AI). Contour plots of the value distributions are shown for threshold, dynamic range, strongest response level (SRL), and monotonicity. For the first three parameters, the *shaded areas* correspond to the approximate lower half of the paramter range for each case. For the monotonicity distributions, the *shaded areas* represent values of moderate to high non-monotonicity. The *thin dashed lines* indicate the orientation of the isofrequency axis. The *thick dashed line* separates dorsal AI (*AId*) from ventral AI (*AIv*), as suggested by a reversal in the overall gradient of integrated excitatory bandwidth. It represents the mean location of the sharpest frequency tuning 10 and 40 dB above minimum threshold and closely approximates the locations of lowest response strength to a broadband transient (Schreiner and Mendelson 1990)

Monotonicity

A major physiological distinction among auditory cortical neurons is the monotonicity or nonmonotonicity of their rate level functions in response to tones (Brugge et al. 1969; Erulkar et al. 1956; Evans and Whitfield 1964; Phillips and Hall 1986; Phillips and Irvine 1981; Phillips et al. 1985). Nonmonotonic units have a distinct maximum in the firing rate for a certain stimulus level. Phillips and Irvine (1981) have reported that 25% of single units in AI had nonmonotonic rate level functions. In another single-unit study of AI (Phillips and Hall 1986), 38.5% of the rate level functions were classified as nonmonotonic.

In the current study, the degree of monotonicity or nonmonotonicity was assessed in a somewhat different way than in the studies mentioned above, by obtaining the slope of the rate level function above the transition point, i.e., after the rapid growth phase of the function. As a result, one derives a continuous measure of the rate of change of that portion of the rate level function. This measure can be estimated without having to rely on the firing rate at the high level end of the rate level function, as is necessary for other measures (Phillips et al. 1985; Imig et al. 1990). The slope measure can also quantify slow positive growths of the high level portion that are not discerned by those measures. Applying the slope measure of monotonicity to a few examples of published single unit rate level functions (Phillips et al. 1985; Phillips and Hall 1986; Phillips 1988) resulted in slopes ranging from +1.56%/dB to −8.7%/dB (N = 16). These re-evaluated single unit rate/level functions had been obtained with tone rise times between 2.5 and 10 ms, thus bracketing the rise time of 3 ms used in this study. Changes in rise times have been shown to influence the slope of the high level portion and, therefore, to alter the degree of nonmonotonicity (Phillips 1988).

In the current multiple-unit study, 60.5% of the locations showed nonmonotonic rate/level functions as judged by a negative slope of the high-level portion, a clear deviation from the pattern seen for single units. This proportion of nonmonotonic locations can be lowered to approximately 20% by using a more strict criterion for manifest nonmonotonicity, for example −1% dB, thereby excluding those moderately negative slope estimates that, at least partially, could be attributed to some random scatter in the course of otherwise saturated rate level functions. However, no unequivocal benchmarks for selecting an appropriate objective criterion for nonmonotonicity were evident.

Another feature of the obtained monotonicity values that is at variance with single-unit data is that the range of the vast majority of multiple-unit slopes was compressed. Values ranged only from +1.5%.dB to −2.5% dB, thus reducing the negative slope range to less than a third of that found for single units.

The main reason for these discrepancies between single- and multiple-unit results is likely related to inherent properties of the multiple-unit recording technique, namely the integration of the firing rate from several units near a given cortical location. In order to observe nonmonotonicities in the multiple-unit recording that are just as large as observed for single units, all units in the recorded cluster have to possess the same degree of nonmonotonicity, the same response threshold, and the same dynamic range. On the other hand, it is less likely that a cluster containing only nonmonotonic units that differ in those other properties would result in a monotonic rate level function. The consequence is that clusters with predominantly nonmonotonic units will also appear as nonmonotonic in the multiple-unit recording, however, with a reduced degree of nonmonotonicity. Clusters that contain only monotonic units will be recorded as monotonic, possibly with a small to moderately large positive slope. The slope of rate level functions for clusters that contain about the same number of monotonic and nonmonotonic units will likely be negative. The magnitude of the negative slope will probably be small and will be determined by the amount of nonmonotonicity as well as by the threshold and dynamic range distributions of the units. In other words, the measure of monotonicity as applied in this study reflects the degree of homogeneity in the rate level functions of the contributing neurons as well as the global nature of the level-dependent behavior of the units at the high and low ends of the monotonicity scale.

Phillips et al. (1985; Fig. 8) demonstrated an example of the spatial distribution of monotonicity at locations in auditory cortex. Of the 58 locations reported for that case, 29.3% showed only cells with monotonic rate level functions, and 48.3% exclusively contained nonmonotonic cells. At the remaining 22.3% locations, monotonic as well as nonmonotonic cells were encountered ("mixed" composition). Combining locations of nonmonotonic cells with those of mixed composition results in 70.6% locations that could show some degree of nonmonotonicity if sampled with a multiple-unit recording. This percentage is in closer agreement with the percentage of nonmonotonic locations (60.5%) found in the present study than with the distribution based on single units that are not topographically localized (see above).

The spatial distribution of the monotonicity measure across the mapped auditory cortices consistently showed aggregates of locations with high nonmonotonicity and usually smaller areas only containing locations with rate level functions of positive slope. In most cases, these areas were separated by locations with slopes near zero, suggesting gradual transitions between predominantly monotonic and predominantly nonmonotonic portions of AI. Indications of spatial segregation of monotonic and nonmonotonic units in AI have been given previously (Reale et al. 1979; Phillips et al. 1985). Phillips and colleagues (1985) also mentioned a "horizontal" segregation of the area of nonmonotonicity from locations with monotonic properties corresponding to the rostrocaudal orientation of the elongated central nonmonotonic area observed in all three cases in this study. In two of the three cases described here, a second, noncontiguous, smaller region of nonmonotonicity was observed that was located dorsal to the main area of nonmonotonicity.

For a functional interpretation of the segregation of areas with high monotonicity and high nonmonotonic-

ity, several aspects have to be considered. The obtained range and spatial pattern of monotonicity values within the isofrequency domain provides further evidence that the processing of signal intensity may be systematically organized and distributed in the primary auditory cortex. Regions of high-level selectivity (strongly nonmonotonic) and regions of low-level selectivity (monotonic) are nonrandomly superimposed over regions with high- or low-level sensitivity, thus providing a number of intensity-specific parameter combinations that are potentially of functional significance.

Another aspect of the monotonicity issue is that a high degree of nonmonotonicity also reflects the involvement of a high degree of inhibitory activity (Suga and Manabe 1982; Phillips and Cynader 1985; Phillips et al. 1985). This covariance has consequences for the appropriate spectral and temporal content of signals that are optimally matched for those locations (Phillips 1988), i.e., it provides constraints for the possible "information-bearing parameters" represented at those locations.

Strongest response level

The stimulus level eliciting the highest firing rate of a unit or cluster response is referred to as "most effective sound pressure level" (Brugge and Merzenich 1969) or "best amplitude" (Suga 1977) and implies some tuning to that level. The measure referred to here as "strongest response level" combines all of the aspects of rate/level functions discussed above – response threshold, dynamic range, and monotonicity – and assigns a single variable. Although the reduction of several properties to a single value represents a simplification of these physiological issues, the procedure is useful since it provides a unified picture that is easily referred to, is expressed in units (dB SPL) that directly relate to well-known physical properties of the signal, and is probably closely related to intensity coding (Pfingst et al. 1977; Suga 1977; Suga and Manabe 1982).

The SRL expresses a condition of maximal excitatory activity of a single unit or a group of neurons. This condition may or may not be of a specific functional significance for a group of neurons whose individual properties may vary (see above). However, as long as the specific properties of local cortical circuitry and the contributions from different types of neurons are not known, this measure should suffice in assessing the global distribution of local states of activity.

A disadvantage that has to be kept in mind in using the measure of SRL is that it does not express the degree of selectivity for that specific level condition. For strongly nonmonotonic units, a SRL is highly significant (Suga and Manabe 1982). For monotonic units, a SRL may also be derived, e.g., as the highest level in units with slightly increasing rate/level functions or as the average level of the saturated portion in a rate/level function. However, the level selectivity in the two latter cases is low and the measure becomes functionally less significant unless it is viewed in combination with other measures that more specifically express the local properties.

Considering that the highest levels used in this study did not usually exceed 85 dB SPL, the range of SRLs observed in this study (10–80 dB SPL) is in reasonable agreement with the range found for best levels for single units in the auditory cortex of the cat (0–65 dB SPL, Imig et al. 1990; or 12–90 dB SPL, as extracted from data published by Phillips 1988; Phillips and Cynader 1985; Phillips and Hall 1986, 1987; and Phillips et al. 1985, 1989), the macaque (15–95 dB SPL, Brugge and Merzenich 1973), and the mustached bat (15–98 dB SPL, Suga 1977; Suga and Manabe 1982). The relatively even distribution of the SRL values across that range in combination with the existence of a significant percentage of monotonic rate/level functions provides evidence for relatively equal signal processing capacities at all intensities in the normal hearing range.

Similar to the other intensity-related measures, SRLs were not randomly distributed across AI but showed aggregates of high, medium, or low SRLs. One major rostrocaudally elongated aggregate of lower SRLs was located approximately in the dorsoventral center of AI oriented orthogonally to the isofrequency domain. The dorsal half of AI was usually dominated by higher SRLs, although smaller clusters of lower SRLs were also observed. In the ventral portion of two maps, SRLs around 60 dB SPL dominated.

The consistent correlation between SRL and threshold (see also Suga and Manabe 1982) as well as monotonicity is reflected in the spatial distribution of the SRL. Superimposing the shaded areas in the threshold and monotonicity contour maps of cases 87–518 and 87–706 provides a good prediction of the distribution of low SRLs. However, the prediction falters in areas that are not highly nonmonotonic and are therefore not particularly level-selective.

The observation that essentially the full range of possible SRLs was encountered in a somewhat systematic fashion along a given isofrequency contour suggests the existence of an ampliotopic representation in cat AI similar to that seen in the mustached bat (Suga and Manabe 1982).

Comparison of spatial characteristics of AI

Frequency organization. The most prominent spatial pattern in the physiological organization of AI is the cochleotopic gradient that unequivocally defines its rostral and caudal extent (e.g., Merzenich et al. 1975; Reale and Imig 1980). The dorsoventral extent of the isofrequency domain is marked by a more or less gradual decline of the precision in the frequency organization toward the presumed boundaries of AI (Merzenich et al. 1975; Middlebrooks and Zook 1984; Reale and Imig 1980; Schreiner and Cynader 1984; Schreiner and Mendelson 1990). Just as in the cytoarchitecture (Rose 1949; Winer 1984) or thalamocortical projection pattern (Andersen et al. 1980; Brandner and Redis 1990; Imig and Morel 1984; Middlebrooks and Zook 1984; Morel and Imig 1987) the CF organization does not provide sharp and unequivocal boundary criteria for the dorsoventral ex-

tent of AI. The distribution of physiological response parameters other than frequency could possibly provide the desired information. However, examination of the maps in this report does not indicate clear parametric criteria that would justify their sole use in determining the dorsoventral boundaries of AI. Most changes in the spatial distribution of parameters are relatively gradual, and in the dorsal or ventral portion of AI, none of the four parameters showed a clear functional demarcation that continuously extended across the rostrocaudal dimension of the middle portion of the ectosylvian sulcus.

Spectral bandwidth. In a previous report (Schreiner and Mendelson 1990), the spatial distributions of the integrative excitatory bandwidth and the response to transient signals had been documented for the same cortices explored in this study. The spatial distribution of the integrated excitatory bandwidth, expressed as Q–10 dB and Q–40 dB of multiple-unit tuning curves, revealed an area of sharp tuning in the dorsoventral center of AI orthogonally oriented to the isofrequency domain (Schreiner and Mendelson 1990) with a more or less gradual decrease of the integrated bandwidth toward the dorsal and ventral boundaries of AI. On the basis of these multiple-unit gradients, as well as the responses of single neurons (Sutter and Schreiner 1991a), AI was divided into a dorsal (AId) and a ventral portion (AIv). Linear regression revealed that response threshold, monotonicity, and SRL showed a more robust correlation with the excitatory bandwidth values in AIv than in AId. This trend is less prominent but still discernible for the correlations with the response to a broadband transient and supports the notion of a physiological distinction between these two regions of AI.

Large responses to a transient signal were predominantly found outside the areas of sharp tuning (Schreiner and Mendelson 1990). A similar reciprocal relationship between nonmonotonicity and large transient responses is suggested by their essentially nonoverlapping spatial distributions (compare Fig. 11 with Fig. 16 in Schreiner and Mendelson 1990). Since units with nonmonotonic responses usually have a narrow frequency tuning (see Table 3) and reflect strong inhibitory influences, the observed relationship between nonmonotonicity and broadband stimulus response strength was to be expected (Phillips et al. 1985). The spatial coincidence between the main area of sharpest frequency tuning and local minima in the distribution of SRL and monotonicity indicate that these parameters are not independently organized in AI.

Binaural interaction. Three of the four described parameter distributions indicate an alignment of at least one of their main spatial aggregates with the CF gradient and, consequently, reflect orthogonality with the isofrequency domain. The main orientational axis of binaural interaction "bands" has also been shown to be orthogonal to the isofrequency domain of AI of the cat (Imig and Adrian 1977; Imig and Brugge 1978; Middlebrooks et al. 1980; Schreiner and Cynader 1984). Although the binaural organization of the maps under study will be described elsewhere, it is useful to say that only the

response threshold showed some degree of correlation with binaurality as determined by simple classification of near-threshold binaural summation or suppression in agreement with observations by Schreiner and Cynader (1984).

Recently, Imig and colleagues (1990) studied a functionally and behaviorally more relevant measure of binaural interaction, that of azimuthal direction selectivity, and reported that the selectivity of neurons in cat AI to the azimuthal direction of a broadband sound source was related to the level selectivity of the neurons. Neurons with strong nonmonotonic rate level functions in response to broadband noise were always highly directional selective and a wide range of SRLs was associated with a particular azimuth. In the same report, the authors stated that locations with similar azimuth and level tuning showed evidence of spatial clustering and that several azimuthal directions were represented within a narrow frequency range across AI.

Deviations between the spatial distribution of low SRL and nonmonotonicity, as reported in this study, support the notion that sharply tuned SRLs, can be found across the whole intensity range. However, highly nonmonotonic locations, i.e.. sharply tuned to a stimulus level, seem to be more prominent in areas with low SRLs, at least for pure tone nonmonotonicity. This may differ from broadband rate/level functions as used in the study by Imig et al. (1990).

Some implications for functional organization of AI

The three illustrated maps were chosen because the distribution of several other response parameters are known for the same cases and . ot because these were the most convincing examples of a systematic representation of sound intensity. The suggestion of a nonrandom distribution of response properties that are involved in the coding of sound intensity, i.e., threshold. SRL, and monotonicity, has been seen in essentially every auditory cortical map derived in our laboratory (e.g.. Merzenich et al. 1991; Raggio et al. 1992; Sutter and Schreiner 1991b). However, the apparent variance from one animal to another in the spatial distribution of intensity-related parameters indeed appears to be larger than that seen for CF (e.g.. Merzenich et al. 1975; Reale and Imig 1980) or bandwidth organization (Schreiner and Mendelson 1990; Sutter and Schreiner 1991a. b).

An important contribution to the idiosyncratic or "noisy" appearance of a potentially ampliotopic signal representation – aside from the influences introduced by the multiple-unit recording method – may have been the fact that the values of the utilized intensity parameters can be modified substantially, e.g., by certain background noise conditions (for review see Phillips 1990) or by certain attention- or learning-induced histochemical response modulations (e.g.. Metherate et al. 1990). Since the current results were obtained without background noise and in the anesthethized animal. it is likely that the cortical representation of intensity parameters under behaviorally more relevant conditions may differ somewhat from those described here.

Following a scheme used in establishing hierarchical organization in visual cortex from ascending and descending projections. Rouiller and colleagues (1991) delineated an overall hierarchy for auditory cortical fields from their interconnections. AI appears to be at the initial level of processing and representation in auditory cortex and, therefore, is likely to provide a more general signal representation based on a number of coding aspects. Consequently, it is possible that the need to represent other signal aspects as well, e.g., binaural or bandwidth, may interfere with a smooth representation of sound intensity in AI. Auditory fields located higher in the hierachy of cortical processing may prove to show more specialized and, consequently, more systematic and spatially smooth representations of certain signal aspects. There is some evidence that the posterior auditory field may contain a more precise map of sound intensity (Phillips and Orman 1984; personal observations) than seen in AI.

Previous physiological and projectional studies provided evidence for a somewhat "patchy" organization of the AI of the cat. In the isofrequency domain, binaural interaction sound localization properties appear to show some banding or clustering of similar response characteristics (Imig and Adrian 1977; Imig and Brugge 1978; Imig et al. 1990; Middlebrooks et al. 1980; Rajan et al. 1990). Intrinsic connections in AI also appeared to be patchy rather than continuous (Matsubara and Phillips 1988; Ojima et al. 1991; Wallace and Bajwa 1991; Wallace et al. 1991a, b). The finding of a rather patchy and overall less gradual spatial distribution of intensity parameters may relate to those anatomical findings and provides further evidence that the isofrequency domain in AI is functionally subdivided and does not represent a uniform representational axis.

References

Andersen RA, Knight PL, Merzenich MM (1980) The thalamocortical and corticothalamic connections of AI, AII and the anterior auditory field (AAF) in the cat: evidence for two largely segregated systems of connections. J Comp Neurol 194:663–701

Brandner S, Redis H (1990) The projection from the medial geniculate to field AI in cat: organization in the isofrequency dimension. J Neurosci 10:50–61

Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. J Neurophysiol 36:1138–1159

Brugge JF, Dubrovsky NA, Aitkin LM, Anderson DJ (1969) Sensitivity of single neurons in auditory cortex of cat to binaural tonal stimulation: effect of varying interaural time and intensity. J Neurophysiol 32:1005–1024

Erulkar SD, Rose JE, Davies PW (1956) Single unit activity in the auditory cortex of the cat. Bull Johns Hopkins Hosp 99:55–86

Evans EF, Whitfield IC (1964) Classification of unit responses in the auditory cortex of the unanaesthetized and unrestrained cat. J Physiol (Lond) 171:476–493

Imig TJ, Adrian HO (1977) Binaural columns in the primary field (AI) of auditory cortex. Brain Res 138:241–257

Imig TJ, Brugge JF (1978) Relationship between binaural interaction columns and commissural connections of the primary auditory field (AI) in the cat. J Comp Neurol 182:637–660

Imig TJ, Morel A (1984) Topographic and cytoarchitectonic organization of thalamic neurons related to their targets in low-, middle-, and high-frequency representations in cat auditory cortex. J Comp Neurol 227:511–539

Imig TJ, Irons WA, Samson FR (1990) Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. J Neurophysiol 63:1448–1466

Jenkins WM, Merzenich MM (1984) Role of cat primary auditory cortex for sound localization behavior. J Neurophysiol 52:819–847

Matsubara JA, Phillips DP (1988) Intracortical connections and their physiological correlates in the primary auditory cortex (AI) of the cat. J Comp Neurol 268:38–48

Mendelson J, Schreiner CE, Grasse K, Sutter ML (1988) Spatial distribution of responses to FM sweeps in cat primary auditory cortex. Assoc Res Otolaryngol Abstr 11:36

Merzenich MM, Knight P, Roth GL (1975) Representation of the cochlea within the primary auditory cortex in the cat. J Neurophysiol 38:231–249

Merzenich MM, Schreiner CE, Recanzone G, Beitel R, Sutter (1991) Topographic organization of cortical field AI in the owl monkey (*Aotus trivirgatus*). Assoc Res Otolaryngol Abstr 14:44

Metherate R, Ashe JH, Weinberger NM (1990) Acetylcholine modifies neuronal acoustic rate-level functions in guinea pig auditory cortex by an action at muscarinic receptors. Synapse 6:364–368

Middlebrooks JC, Pettigrew JD (1981) Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. J Neurosci 1:107–120

Middlebrooks JC, Zook JM (1983) Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. J Neurosci 3:203–224

Middlebrooks JC, Dykes RW, Merzenich MM (1980) Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. Brain Res 181:31–48

Morel A, Imig TJ (1987) Thalamic projections to fields A, AI, P, and VP in the cat auditory cortex. J Comp Neurol 265:119–144

Ojima H, Honda CN, Jones EG (1991) Patterns of axon collateralization of identified supragranular pyramidal neurons in the cat auditory cortex. Cerebral Cortex 1:80–94

Pfingst BE, O'Connor TA, Miller JM (1977) Single cell activity in the awake monkey cortex: intensity coding. Trans Am Acad Ophtalmol Otularyngol 84:217–222

Phillips DP (1988) Effects of tone-pulse rise time on rate-level functions of cat auditory cortex neurons: excitatory and inhibitory processes shaping responses to tone onset. J Neurophysiol 59:1524–1539

Phillips DP (1990) Neural representation of sound amplitude in the auditory cortex: effects of noise amplitude. Behav Brain Res 37:197–214

Phillips DP, Cynader MS (1985) Some neural mechanisms in the cat's auditory cortex underlying sensitivity to combined tone and wide-spectrum noise stimuli. Hear Res 18:87–102

Phillips DP, Irvine DRF (1981) Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. J Neurophysiol 45:48–58

Phillips DP, Hall SE (1986) Spike-rate intensity functions of cat cortical neurons studied with combined tone-noise stimuli. J Acoust Soc Am 80:177–187

Phillips DP, Hall SE (1987) Response of single neurons in cat auditory cortex to time-varying stimuli: linear amplitude modulations. Exp Brain Res 67:479–492

Phillips DP, Orman SS (1984) Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. J Neurophysiol 51:147–163

Phillips DP, Hall SE, Hollett JL (1989) Repetition rate and signal level effects on neuronal responses to brief tone pulses in cat auditory cortex. J Acoust Soc Am 85:2537–2549

Phillips DP, Orman SS, Musicant AD, Wilson GF (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. Hear Res 18:73–86

Raggio ME, Schreiner CE, Merzenich MM (1992) Correspondence of functional topographies in cat primary auditory cortex for acoustic and electric cochlear stimulation. Soc Neurosci Abstr 18:382

Rajan R, Aitkin LM, Irvine DRF (1990) Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. J Neurophysiol 64:888–902

Reale RA, Imig TJ (1980) Tonotopic organization in cat auditory cortex. J Comp Neurol 192:265–292

Reale RA, Kettner RE (1986) Topography of binaural organization in primary auditory cortex of the cat: effects of changing interaural intensity. J Neurophysiol 56:663–682

Reale RA, Imig TJ, Sinex DG (1979) Rate-intensity functions of single neurons located within binaural suppression columns of cat primary auditory cortex. Soc Neurosci Abstr 5:29

Rose JE (1949) The cellular structure of the auditory area of the cat. J Comp Neurol 91:409–439

Rouiller EM, Simm GM, Villa AEP, De Ribaupierrre Y, De Ribaupierre F (1991) Auditory corticocortical interconnections in the cat: evidence for parallel and hierarchical arrangement of the auditory coortical areas. Exp Brain Res 86:483–505

Schreiner CE, Cynader MS (1984) Basic functional organization of second auditory cortical field (AII) of the cat. J Neurophysiol 51:1284–1305

Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. J Neurophysiol 64:1442–1459

Schreiner CE, JR Mendelson Grasse K, Sutter M (1988) Spatial distribution of basic response properties in cat primary auditory cortex. Assoc Res Otolaryngol Abstr 11:36

Shamma SA, Fleshman JW (1990) Spectral orientation columns in the primary auditory cortex. Assoc Res Otolaryngol Abstr 13:222

Suga N (1977) Amplitude-spectrum representation in the Doppler-shifted processing area of the auditory cortex of the mustached bat. Science 196:64–67

Suga N (1984) The extent to which biosonar information is represented in the bat auditory cortex. In: Edelman GM, Gall WE, Cowan WM (eds) Dynamic aspects of neocortical function Wiley, New York, pp 315–373

Suga N (1988) Auditory neuroethology and speech processing: complex-sound processing by combination-sensitive neurons. In: Edelman GM, Gall WE, Cowan WM, (eds) Auditory function, neurobiological bases of hearing. Wiley, New York pp 679–720

Suga N, Manabe T (1982) Neural basis of amplitude-spectrum representation in the auditory cortex of the mustached bat. J Neurophysiol 47:225–255

Suga N, Tsuzuki K (1985) Inhibition and level-tolerant frequency tuning in the auditory cortex of the mustached bat. J Neurophysiol 53:1109–1145

Suga N, Niwa H, Taniguchi I, Margoliash D (1987) The personalized auditory cortex of the mustached bat: adaptation for echolocation. J Neurophysiol 58:643–654

Sutter ML, Schreiner CE (1991a) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. J Neurophysiol 65:1207–1226

Sutter ML, Schreiner CE (1991b) Topography of intensity encoding by single neurons in cat primary auditory cortex. Soc Neurosci Abstr 17:1484

Tunturi AR (1952) A difference in the representation of auditory signals for the left and right ears in the iso-frequency contours of the right middle ectosylvian and auditory cortex of the dog. Am J Physiol 168:712–727

Wallace MN, Bajwa S (1991) Patchy intrinsic connections of the ferret primary auditory cortex. Neuroreport 2:417–420

Wallace MN, Kitzes LM, Jones EG (1991a) Intrinsic inter- and intralaminar connections and their relationship to the tonotopic map in cat primary auditory cortex. Exp Brain Res 86:527–544

Wallace MN, Kitzes LM, Jones EG (1991b) Chemoarchitectonic organization of the cat primary auditory cortex. Exp Brain Res 86:518–526

Winer JA (1984) Anatomy of layer IV in cat primary auditory cortex (AI). J Comp Neurol 224:535–567

Woolsey CN, Walzl EM (1942) Topical projection of nerve fibers from local regions of the cochlea in the cerebral cortex of the cat. Bull Johns Hopkins Hosp 71:315–344

Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps

Julie R. Mendelson[1,*], Christoph E. Schreiner[1], Mitchell L. Sutter[1,**], Keith L. Grasse[2]

[1] Coleman Laboratory, Department of Otolaryngology, University of California at San Francisco, San Francisco, CA 94143-0732, USA
[2] Department of Psychology, York University, North York, Ontario, Canada M3J 1P3

Abstract. The spatial distribution of neuronal responses to frequency-modulated (FM) sweeps was mapped with microelectrodes in the primary auditory cortex (AI) of barbiturate-anesthetized cats. Increasing and decreasing FM sweeps (upward- and downward-directed FM sweeps, respectively) covering a range of 0.25–64.0 kHz were presented at three different rates of frequency change over time (i.e., sweep speed). Using multiunit recordings, the high-frequency domain (between 3.2 and 26.3 kHz) of AI was mapped over most of its dorsoventral extent (as determined by the distribution of the excitatory bandwidth, Q_{10dB}) for all six cases studied. The spatial distributions of the preferred sweep speed and the preferred sweep direction were determined for each case. Neuronal responses for frequency sweeps of different speeds appeared to be systematically distributed along the dorsoventral axis of AI. In the dorsal region, cortical cells typically responded best to fast and/or medium FM sweeps, followed more ventrally by cells that responded best to medium – then slow –, then medium-speed FM sweeps. In the more ventral aspect of AI (which in some cases may also have included cells located in the dorsal region of the second auditory field, AII), neurons generally preferred fast FM sweeps. However, a comparison of maps from different animals showed that there was more variability in the distribution of preferred speed responses in the ventral region of the cortex. The directional preference of units for FM sweeps was determined for the sweep speed producing the strongest response. Direction selectivity appeared to be nonrandomly distributed along the dorsoventral axis of AI. In general, units that responded best to upward-directed FM sweeps were located in the more dorsal and ventral aspects of AI while units that responded best to downward-directed FM sweeps were usually located in the mid-region of AI. Direction selectivity was also determined for multiunit responses at each of the three FM sweep speeds. In general, there was a relatively close agreement between the spatial distributions of direction selectivity determined for the strongest response with those calculated for the fast and medium speeds. The spatial distribution of direction selectivity determined for slow FM sweeps deviated somewhat from that determined for the strongest response. Near the dorsoventral center of the mapped areas, the distribution of units that responded best to downward sweeps tended to overlay the distribution of units that responded best to slow speeds, suggesting some spatial covariance of the two parameters. However, when the analysis was extended over the entire region of cortex examined in this study, the point-by-point correlation between preferred speed and direction selectivity was not statistically significant. In addition, when neural responses obtained from the dorsal and ventral subregions were analyzed separately, no significant correlation was observed between these two response parameters. This suggests that, for a given cortical location, the response properties of direction selectivity and preferred speed are derived from distinct neural processing mechanisms. Significant observations were also made between preferred FM sweep speed and excitatory bandwidth (i.e., Q_{10dB} and Q_{40dB}) such that units that responded best to slower FM speeds also seemed to have higher Q_{10dB} and Q_{40dB} (i.e., were narrowly tuned) and vice versa. In addition, units that responded well to a broadband transient stimulus in general preferred faster FM sweeps and vice versa. Although these correlations were significant across the entire dorsoventral extent of AI investigated in this study, they were stronger for responses in the dorsal subregion of AI. For direction selectivity, statistically significant correlations with these response parameters were observed more often in the dorsal than the ventral regions of AI. The apparent spatial segregation of neuronal responses to different FM sweep speeds and sweep direc-

Present address: *Division of Life Sciences, University of Toronto, Scarborough Campus, 1265 Military Trail, Scarborough, Ontario, Canada M1C 1A4

Present address: **Dept. of Organismal Biology, University of Chicago, 1025 E. 57th St., Chicago, IL 60637, USA

Correspondence to: J.R. Mendelson, Division of Life Sciences, University of Toronto, Scarborough Campus, 1265 Military Trail, Scarborough, Ontario, Canada M1C 1A4

tions distributed along the isofrequency domain of AI suggests that the global aspects of cortical function are compatible with psychophysically derived notions of parallel streams of processing for different aspects of FM signals.

Key words: Frequency-modulated sweep – Direction and speed selectivity – Primary auditory cortex – Topographical organization – Cat

Introduction

Naturally occurring auditory signals are characterized by their modulation of amplitude and frequency composition over time. These types of modulation are evident in various aspects of communication signals as well as in sound sources that move relative to an observer. Relevant stimuli for studying neuronal responses to these dynamic acoustical properties include amplitude-modulated (AM) and frequency-modulated (FM) signals. It has been shown in different stages of the auditory pathway that neurons are responsive to AM and FM signals. For AM stimuli, the range of modulation frequencies that produce excitatory responses in neurons decreases along the ascending auditory pathway, indicating the processing of a more limited range of temporal events in higher auditory structures (Rees and Moller 1983; Schreiner and Urbas 1986, 1988; Schreiner and Langner 1988). Ascending in the auditory pathway from the periphery to the cortex, selectivity of neurons for certain features of FM stimuli, such as the direction (i.e, upward-directed: changing from a low to a high frequency, or vice versa) or speed of frequency sweeps (i.e., rate of change of frequency over time), appears to increase (Suga 1965a; Whitfield and Evans 1965; Watanabe 1972; Moller 1974; Nomoto 1980; Shore and Nuttall 1985; Poon et al. 1991; Mendelson and Grasse 1992). At the level of the primary auditory cortex (AI) of cats, the majority of neurons are selective for a particular direction and/or speed of FM sweep (Mendelson and Cynader 1985; Mendelson and Grasse 1992). For example, Mendelson and colleagues have shown in their single unit investigations of AI that, of those cells that are direction selective, the majority prefer downward-directed as opposed to upward-directed FM sweeps (Mendelson and Cynader 1985; Mendelson and Grasse 1992). They also observed that the majority of cortical cells preferred relatively fast speeds (0.4–0.8 kHz ms) of FM sweeps (Mendelson and Cynader 1985; Mendelson and Grasse 1992). In addition, they found that the preferred FM speed and the degree of direction selectivity varied considerably from neuron to neuron.

A pertinent question regarding FM and AM stimulus responses is their functional organization within the auditory cortex. Topographical functional organization appears to be a general feature of neocortical specialization (Mountcastle 1957; Hubel and Wiesel 1965). For example, in the auditory cortex of the bat, the interpretation of functional specialization has been greatly enhanced by understanding the underlying topographical organization (Suga 1965a; O'Neill and Suga 1982; Tsuzuki and Suga 1988; Edamatsu et al. 1989). Suga and his colleagues (Suga 1965a, 1986; O'Neill and Suga 1982; Tsuzuki and Suga 1988; Edamatsu et al. 1989) have shown that neuronal responses to particular aspects of FM stimuli, a crucial component of the bat's biosonar system, are topographically organized in the auditory cortex. In fact, they have identified an entire cortical area devoted exclusively to the processing of FM signals.

Until recently, only two response features were known to be systematically distributed in the auditory cortex of other mammals such as the cat. These features are the characteristic frequency (CF), which is distributed along the rostrocaudal axis (Woolsey and Walzl 1942; Merzenich et al. 1975; Reale and Imig 1980), and alternating bands or patches of binaural facilitatory (EE) and binaural inhibitory (EI) cells, distributed along the dorsoventral extent of AI (Imig and Adrian 1977; Middlebrooks et al. 1980). More recently, Reale and Kettner (1986) extended these results by showing that cells located in the ventral region of cat AI contain both facilitatory and inhibitory response characteristics, depending upon the stimulus intensity presented. Within the last few years, evidence has begun to accumulate which shows that other functional response properties are systematically distributed in AI. For example, Schreiner and Mendelson (1990) have shown that "integrated excitatory bandwidth" (i.e., Q_{10dB} and Q_{40dB}) and responses to broadband transients are topographically arranged within AI. The principal gradient of the integrated excitatory bandwidth was found to be oriented orthogonally to the cochleotopic gradient. Other response parameters such as inhibitory sidebands (Sutter and Schreiner 1991), stimulus intensity responses and monotonicity (Schreiner et al. 1992) have also been found to be systematically distributed in the auditory cortex. In addition, certain properties of sound localization cues have been shown to be systematically distributed in the auditory cortex (Imig et al. 1990; Rajan et al. 1990).

While we now have a greater understanding of the spatial distribution of many discrete cortical response properties, we do not know if or how responses to dynamic signal features are distributed within the auditory cortex. Given that Mendelson and colleagues (Mendelson and Cynader 1985; Mendelson and Grasse 1992) have shown that the majority of auditory cortical single units are selective for the direction and/or speed of FM sweep, we addressed the question of whether neurons with similar FM response properties are randomly distributed across the cortex, or whether FM responses show a systematic spatial organization across the primary auditory cortical field. Specifically, we explored neuronal responses to the *direction* and *speed* of FM sweeps over a wide cortical area to see if the regional distribution of these responses reflects an organized representation of these stimulus parameters in AI.

Materials and methods

Preparation

A more detailed description of the preparation and recording methods has been presented in a previous paper in this series (Schreiner and Mendelson 1990). Briefly, experiments were conducted on six adult cats shown otoscopically to have clean ears. Animals were initially tranquilized with ketamine hydrochloride (10 mg/kg) and acetylpromazine maleate (0.1 mg/kg) intramuscularly to allow for venous cannulation. Animals were then given an i.v. injection of pentobarbital sodium (Nembutal, 30 mg/kg) followed by intramuscular injections of atropine (1 mg) and dexamethasone sodium phosphate (0.14 mg/kg). Animals were maintained at a surgical level of anesthesia throughout the duration of the experiment with an i.v. infusion of Nembutal (2 mg/kg per h) in lactated Ringer solution (infusion volume 3.5 ml/h). Body temperature was maintained at 37.5°C.

Both pinnae were reflected and the auditory meatuses transected to allow for placement of the earphones within 3 mm of the tympanic membrane. The temporal muscle on the right side was retracted and a craniotomy made to expose the right middle ectosylvian gyrus. All wound margins and pressure points were infiltrated with a long-acting local anesthetic (bupivacaine hydrochloride 2.5%). The dura mater was resected and the underlying brain was covered with silicone oil. A photograph of the surface vasculature was used to record the locations of the electrode penetrations.

Stimulation and recording

The animal was placed in a head holder that permitted unobstructed access to the auditory meatuses. The animal was located in an electrically shielded, sound-attenuating chamber (IAC). Insulated tungsten electrodes were introduced into the cortex orthogonal to the surface (as viewed through a Zeiss operating microscope) by means of a remote-controlled stepping motor microdrive. For most cases, single-unit and unit cluster responses were examined at one location per electrode penetration at a depth between 600 and 1000 μm below the cortical surface, corresponding roughly to layers III and IV. However, for a small sample of penetrations, single units were recorded at a number of different depths (ranging from 500 to 1800 μm). The level discriminator for spike selection was set at 50–100% above the noise level of the recording. The number of units per recording location was estimated to be between 2 and 8. Stimulus-evoked action potential event times were collected and stored on-line by a PDP 11/73 computer.

Cats were initially stimulated with tone bursts (50 ms duration with a 3-ms rise/fall time) presented in a pseudorandom sequence of frequency-intensity combinations selected from 15 intensity levels (5-dB steps) and 45 frequency values (geometrically centred around the CF) as well as with broadband transient signals (see Schreiner and Mendelson 1990). Following this, cells were tested with exponential FM sweeps extending from 0.25 to 64.0 kHz (upward-directed) or from 64.0 to 0.25 kHz (downward-directed) generated by a computer-triggered function generator (Wavetek 185). Exponential FM sweeps were used in this study for a number of reasons. First, the broad FM sweep range ensures that the signal starts and ends outside the excitatory and inhibitory response areas of the neurons under study, thus providing a uniform stimulation paradigm. Second, exponential FM sweeps bear a close resemblance to the frequency organization of the basilar membrane resulting in similar FM sweep parameters across the range of frequencies examined without the need to adjust the speed according to the characteristic frequency (CF) of each neuron. Rates of frequency change (i.e., sweep speeds) between 1 and 110 octaves per second were used. However, for the majority of the multiple-unit recordings, only three sweep rates were used: 110 octaves per second, 33 octaves per second, and 10 octaves per second. The corresponding sweep durations for these FM sweep speeds were 72, 240, and 800 ms, respec-

tively. The presentation rate of either an upward- or downward-directed FM sweep was 1 sweep per second and was independent of the sweep speed. The sweep speed was held constant within each FM sweep presentation, but was varied from one set of trials to the next. Each FM sweep was preceded by a 200-ms interval containing the upper or lower starting frequency of the sweep (both of these frequencies were beyond the average hearing range of the cat by at least 1.3–2.5 octaves from the CF) in order to avoid the risk of contaminating the data with any equipment related broadband transient click at the beginning of each FM sweep presentation. For the FM sweeps whose speeds were 110 and 33 octaves per second the FM sweeps were followed by a 628- or 560-ms interval of the upper or lower final frequency of the FM sweep, respectively. The onset of each FM sweep stimulus was triggered by the computer.

Both the pure tone and FM sweep stimuli were attenuated (Hewlett-Packard) and passed through transducers (STAX 54) attached to sound delivery tubes sealed into the auditory meatuses. The system was calibrated for pure tones with a sound-level meter (Bruel and Kjaer 2209). The transfer function of the sound delivery system showed a roll-off of 10 dB/octave above 14 kHz, and additional deviations of up to ± 6 dB were possible (see Fig. 1 in Schreiner and Mendelson 1990).

Recording procedure

Initially, the CF, response threshold, and frequency response area of a neuron or group of neurons were determined. Among other parameters, the excitatory bandwidth at 10 dB and 40 dB above minimum threshold (expressed as Q_{10dB} and Q_{40dB}, respectively) and response to a broadband transient stimulus were also determined (see Schreiner and Mendelson 1990). Following these characterizations, FM sweeps were typically presented monaurally to the contralateral ear. For the small sample of cells encountered which, by conventional standards, would be classified as predominantly binaural cells, FM sweeps were presented binaurally.

FM sweeps were presented at a stimulus level approximately 40 dB above the minimum threshold. This stimulus level was selected for a number of reasons. First, we wanted to include the majority of neurons at a given recording site. If FM sweeps were presented at 10 or 20 dB above threshold, then units with high thresholds may have been inadvertently excluded from the sample. Second, by presenting FM sweeps at 40 dB above threshold, more units would be stimulated maximally. Finally, while it has been shown in single-unit studies that nonmonotonic units can exhibit a rate decrease with an increase in stimulus intensity (Phillips and Irvine 1981), we felt that the possibility of recording weaker responses in some nonmonotonic cells was preferable to the possibility of not recording any responses from cells with higher thresholds.

Each of the six FM sweep conditions (two directions at three speeds) was presented 40 times, and poststimulus time histograms (PSTHs) were constructed. The number of spikes within a 10-ms window around the maximum response for each condition was used as a measure of response strength for each stimulus condition.

With the exception of one experiment, a cortical surface area of approximately 3×5 mm was examined with 80–150 electrode penetrations per map. In one case, the cortical area sampled was 3×2.2 mm. The recording strategy was to map the dorsoventral domain of AI corresponding to the orientation of the isofrequency axis. In several cases, the most dorsal aspects of the second auditory field (AII) were included. The range of CFs investigated across all animals extended from 3.2 to 26.3 kHz.

Data analysis

Frequency response areas derived from the 675 different frequency-intensity combinations were determined for every unit cluster. From this we were able to derive the integrated excitatory bandwidth at 10 and 40 dB (Q_{10dB} and Q_{40dB}, respectively). Following this, two

methods were used to analyze the speed sensitivity of neurons at a given cortical location. The first type of analysis took into account the strength of responses for all three FM speeds tested and provided a non-quantal measure of preferred speed, independent of the total firing rate at a given cortical location. For this type of analysis, the following equation was used:

$$(BR_S + (2 \times BR_M) + (3 \times BR_F))/(BR_S + BR_M + BR_F)$$

where BR is the best response (i.e., the highest firing rate) of the two directions at each of the three FM sweep rates (S, slow; M, medium; F, fast). This resulted in a number between 1 and 3 where 1 represented an exclusive response to the slow speed and 3 an exclusive response to the fast speed. Numbers close to 2 either represented a response to the medium speed or indicated a response with equal response to all three speeds, i.e., a response without speed selectivity. This potential ambiguity affected an average of 10% of all cortical locations examined. However, the second type of analysis for FM speed selectivity described below, helped to further distinguish whether a particular response was for medium FM sweep speeds or was non-speed selective

The second type of analysis provided a discrete measure of the preferred speed as compared to the continuous, nondiscrete mea-

sure described above. This second method involved selecting the speed condition that elicited at least 25% more spikes than either of the other two FM speeds and assigning it the number 1 if that condition was the slow FM sweep, the number 2 if it was a medium sweep, and the number 3 if it was a fast FM sweep. For those cases where the response difference between slow and medium or medium and fast was less than 25%, but each exceeded that of the remaining sweep speed by more than 25%, the numbers 1.5 and 2.5 were assigned, respectively. If none of these criteria were met, the location was classified as nonselective for FM speeds.

The direction selectivity (DS) of a neuron or neuron cluster was determined by comparing its response to upward- and downward-directed FM sweeps at its preferred speed. The degree of direction selectivity was quantified using the following equation:

$$DS = RU - RD/(RU + RD)$$

where RU the firing rate elicited by upward-directed FM sweeps and RD to downward-directed FM sweeps. Values ranged from +1 (responding exclusively to upward-directed FM sweeps) to −1 (responding exclusively to downward-directed FM sweeps). This measure of direction selectivity was insensitive to the units' overall firing rate and thus, was not confounded by rate changes due to the speed
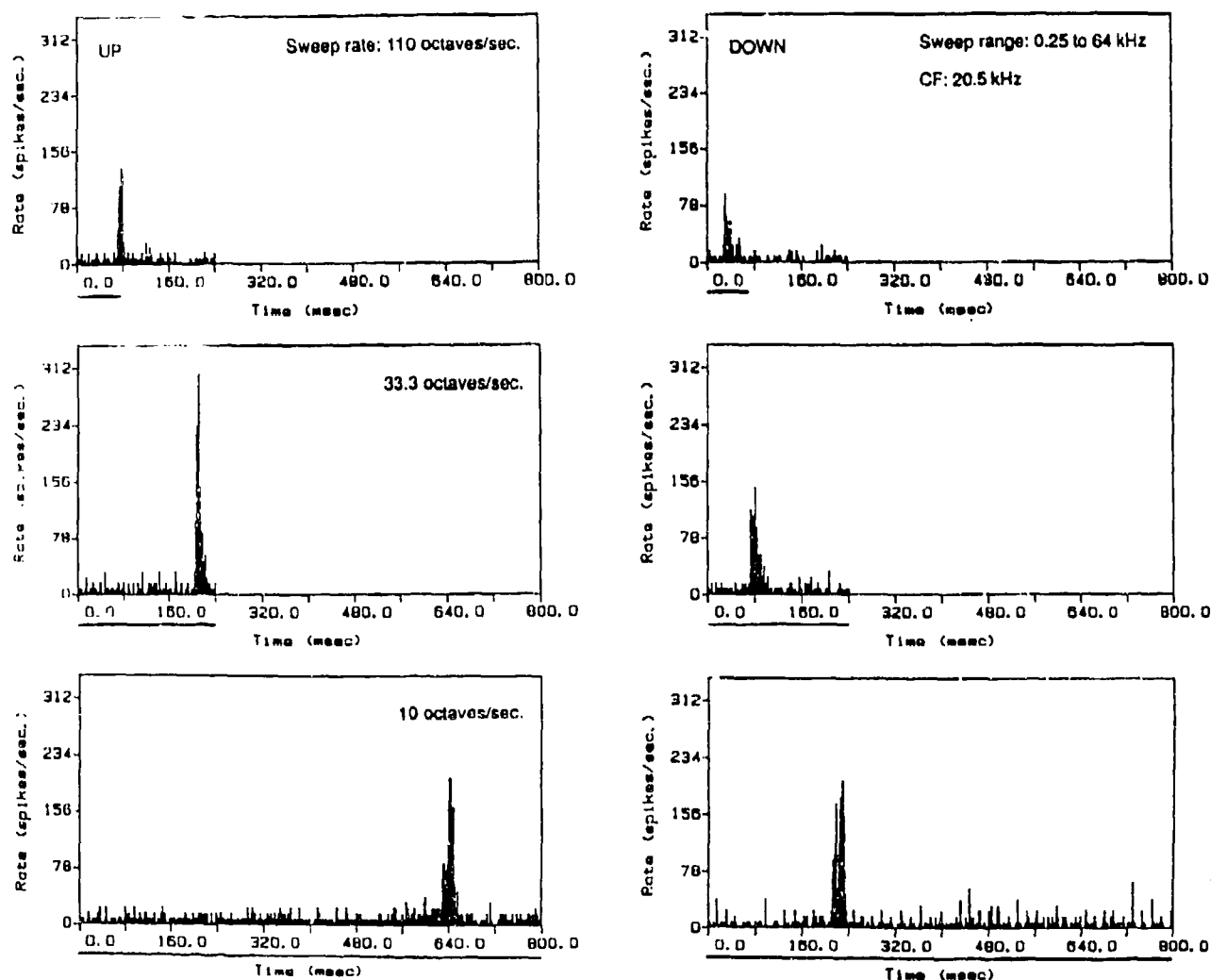
Fig. 1. An example of a multiunit response to frequency-modulated (FM) sweeps. The panels on the *left* illustrate the responses to upward-directed FM sweeps, while those on the *right* are to downward-directed FM sweeps. The responses to the fastest FM sweeps are shown in the *top two panels*, the responses to the medium FM sweeps are in the *middle*, and the responses to the slowest are shown in the *bottom two panels*. The *bar* under each *abscissa* depicts the duration of the FM sweep for each of the three speeds. This unit cluster responded most vigorously to upward-directed FM sweeps at a speed of 33.3 octaves per second

of FM sweep. For example, one might pre !.. ' that the firing rate of a unit cluster to a slow FM sweep could be greater than the firing rate of a unit cluster to a fast FM sweep because the multi-unit would be stimulated in its excitatory response area for a longer period of time with slow FM sweeps than if it were stimulated with a faster FM sweep. Unit responses were considered to be direction selective if their index value was equal to or greater than ±0.15, indicating approximately 35% more spikes for one direction than the opposite direction.

The spatial distribution of the response properties are presented as pseudo three-dimensional projections. These types of plots were generated from the original data by a software package (Golden Software). The interpolation algorithm used (inverse distance squared) inherently applied some minimal spatial smoothing to the data, resulting mainly in a slight compression of the encountered functional range and a local averaging of closely spaced adjacent points (see Schreiner and Mendelson 1990). This method emphasizes topographical organizations with low spatial frequency and de-emphasizes steep local gradients or variabilities. Examples are given that also show the raw data, thus allowing a direct evaluation of the appropriateness of the pseudo three-dimensional representation of the data.

Results

The present results are based on 506 single- and multi-unit responses encountered in closely spaced electrode penetrations over a wide region of AI in six animals. The distance between each penetration varied from 150 to 350 μm. With the exception of one case (87 420), the dorsoventral extent of cortex mapped in these experiments was approximately 5 mm. Within each map, a large portion (>75%) of the tonotopic map indicative of AI was observed (Merzenich et al. 1975; Reale and Imig 1980). All recording sites examined contained responses to pure tones and to at least one of the FM sweep direction and/or speed conditions. The results based on the integrated excitatory bandwidth and responses to broadband transient stimuli are described in detail by Schreiner and Mendelson (1990) and thus will only be summarized here at the end of the Results.

Figure 1 illustrates a typical set of responses from a multiunit cluster at a single cortical site to the six FM sweep conditions examined in this study. The three panels on the left show multiunit responses to upward-directed FM sweeps, while the three panels on the right show responses to downward-directed FM sweeps. Figure 1 also illustrates how cortical responses are affected by different FM sweep speeds: the top two panels show responses to fast sweeps (110 octaves per second), the middle panels show responses to medium sweeps (33 octaves per second), and the bottom panels show responses to slow sweeps (10 octaves per second). As can be seen, cells at this cortical site responded best to upward-directed FM sweeps presented at the medium speed. Generally, only one dominant response maximum in the PSTH was observed during each FM sweep presentation. Occasionally, a second smaller response was present. However, the relative timing of secondary maxima (relative to the dominant response maximum) was independent of the sweep rate used. This suggests a rebound effect rather than a response to a second excitatory frequency region or an amplitude modulation effect.
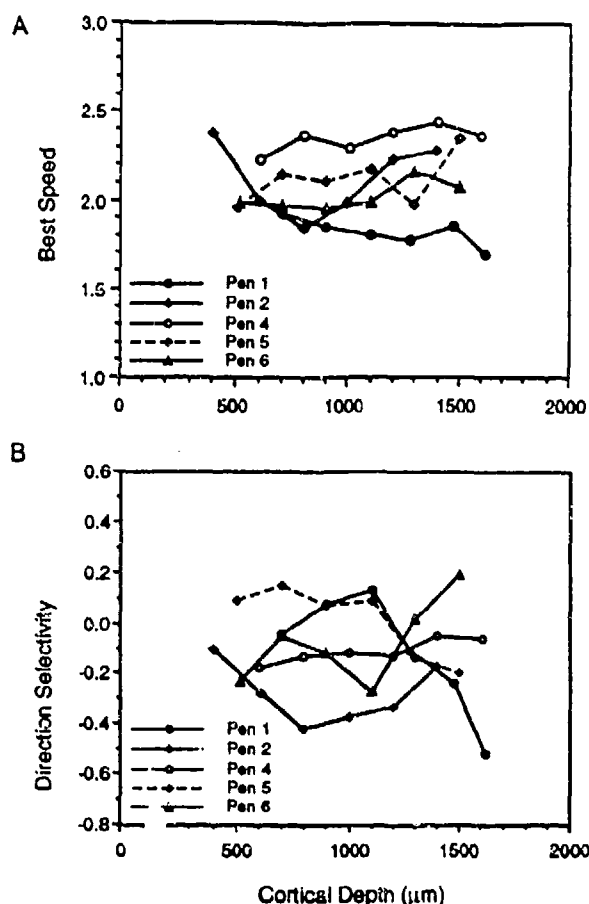
Fig. 2. Dependence of preferred FM speed selectivity (A) and direction selectivity (B) on cortical recording depth. The response parameters for the same five penetrations are plotted as a function of recording depth (radial penetrations). No statistically significant correlation between recording depth and preferred FM sweep speed or direction selectivity was observed. Pen 1, penetration 1; Pen 2, penetration 2; etc

While most multiunit cortical responses described here were recorded at a single depth per electrode penetration, for one case we examined single unit responses at several depths within a single penetration. Figure 2 shows examples of preferred FM sweep speed responses (Fig. 2A) and direction-selective responses (Fig. 2B) as a function of cortical depth for five different electrode penetrations. As illustrated, neither the preferred speed nor the preferred direction appeared to change significantly as a function of cortical depth, especially for responses encountered more than 650 μm from the cortical surface.

As described in Materials and methods, responses to FM sweeps were obtained using a stimulus level of 40 dB above response threshold. For a small sample of multi-units ($n = 10$), however, we examined the responses to FM sweeps presented over a range of stimulus levels to ascertain the possible effects of this variable on preferred FM speed (Fig. 3A) and preferred direction (Fig. 3B). The linear regression analysis for stimulus level and preferred FM speed was statistically significant. However, the slope of the line indicates only a small shift toward preferences
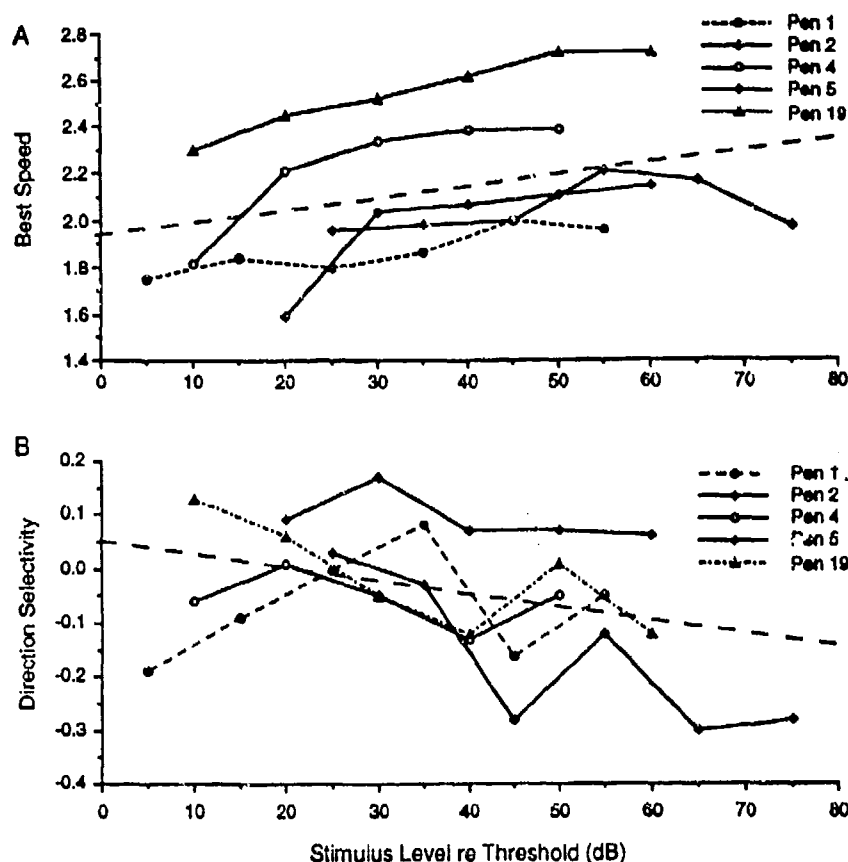
Fig. 3A,B. Dependence of preferred FM speed and direction selectivity on stimulus level of FM sweep. A The preferred FM sweep speed is plotted as a function of stimulus level re minimum threshold (0 dB). Responses from five recording locations from the middle cortical laminae are shown. The *dashed line* is the result of a linear regression analysis $[y = 0.07 \times -0.33; r = 0.76; P = 0.0001$ (*F*-test)]. B The linear regression analysis for direction selectivity as a function of stimulus level is indicated by the *dashed line* $[y = -0.02 \times +0.08; r = 0.41; P = 0.03$ (*F*-test)]. *Pen 1*, penetration 1; *Pen 2* penetration 2; etc

for faster FM sweeps at higher stimulus levels. For the effects of stimulus level on direction selective responses, the linear regression analysis was also statistically significant (Fig. 3B). However, as with preferred FM speed and stimulus level, the shallow slope of the line suggests that there is a minimal change from upward- to downward-directed FM sweeps as stimulus level is increased.

Representation of preferred speed

Table 1 provides the distribution of preferred FM sweep rates for each individual case, as well as for all cases combined. In general, the largest percentage of recording sites (39.8%) was marked by responses preferring fast

FM sweep rates. In all but one case (86-697), preferences for fast FM sweeps were encountered at least twice as often as preferences for medium or slow FM sweeps. The most striking anisotropy for fast versus medium and slow sweep speeds was observed in case 87-129 where 61.7% of the responses encountered preferred fast FM sweep rates. Excluding those sites in which responses were strongly driven by more than one rate of FM sweep, the second largest percentage of observed responses (16.7%) preferred medium-speed FM sweeps, followed by yet a smaller percentage (9.3%) that preferred slow FM sweeps. At approximately 10% of the cortical locations examined, recorded responses exhibited no preference for the three FM speeds studied.

In order to demonstrate that the FM sweep speeds

Table 1. Distribution of preferred frequency-modulated sweep rates

Cat	Number of sweeps	Slow (%)	Slow/ medium (%)	Medium (%)	Medium/ fast (%)	Fast (%)	No preference (%)
86-697	88	5.7	9.1	27.3	20.3	27.3	10.2
87-420	63	14.3	1.6	15.9	20.6	41.3	6.3
87-518	81	12.4	4.9	13.6	16.0	42.0	11.1
87-706	83	3.6	8.4	20.5	19.3	41.0	7.2
87-129	81	6.2	2.5	6.2	11.1	61.7	12.3
87-001	96	14.6	8.3	15.6	22.9	29.2	9.4
Total	492	9.3	1..	16.7	18.5	39.8	9.6
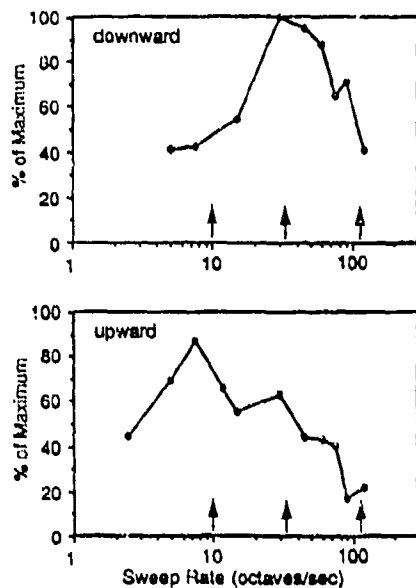
Fig. 4. Rate tuning profile of a single-unit to a range of FM sweep speeds. This neuron is representative of a small sample of single-units presented with a range of sweep speeds, both slower and faster than the three used in the study. The response is expressed as the percentage of the maximum that occurred for downward- directed FM sweeps modulating at a rate of 33 octaves per second. The *arrows* indicate the three sweep speeds used in the study. In general the responses to the additional FM sweep speed conditions provided a gradual transition between the three speeds used to generate the spatial distribution

Fig. 5A,B. Scatter diagrams of preferred FM sweep rate as a function of characteristic frequency (*CF*). All of the preferred speed responses are plotted as a function of CF for two cases. In general, preferred speed appeared to be independent of CF. The *number 1* represents a preference for slow speeds, the *number 2* for medium speeds, and *3* for fast speeds

used in the present study (a) were within the range of speeds that cortical neurons can respond to and (b) sampled the appropriate range of speeds with sufficient density, a wider range of FM speeds was presented to a sample of single cortical units ($n = 10$). This expanded range of FM speed conditions included slower and faster rates than those used in the mapping studies. These FM rates were spaced widely enough to cover a broad ra··· of speeds yet narrowly enough to detect prominent peaks in the speed profile. Figure 4 shows a representative example of the rate tuning profile of a single unit to this broader range of FM sweep speeds. The response is expressed as a percentage of the maximum response. The arrows indicate those three speed conditions that were used to determine the spatial distribution of FM sweep responses in the present study. As illustrated, this neuron responded best to downward-directed FM sweeps changing at a rate of 33 octaves per second; i.e., the medium FM sweep speed condition used throughout the mapping study. In general, FM speed profiles were unimodal, fairly broadly tuned, and showed smooth transitions in the response strength.

Figure 5 shows the distribution of preferred speeds of FM sweep as a function of CF for two cases in which a narrow and a broad range of CFs were sampled, respectively. In both cases, preferred speed did not seem to depend upon CF. Further, there was a uniform distribution of values over the range between 1.5 and 2.5, although some clustering of values around 2.0 was discernible.

In most cases, there appeared to be an organized, non-random spatial distribution of preferred speed of FM sweep responses in AI. Figure 6 illustrates an example of a spatially reconstructed distribution of preferred FM speeds across a high-frequency region of AI. The cortical sites recorded throughout this region were neural responses with CFs between 9.6 and 12.4 kHz. In the pseudo three-dimensional projections, the elevation of the surface of the plot (along the Z-axis) represents the preferred speed sampled at a given recording location. Thus, low elevations represent cortical locations containing preferences for slow FM sweeps and higher elevations indicate cortical regions in which preferences for fast FM sweeps were observed. The dotted line indicates a preference for medium FM speeds (denoted by the number 2) and has been included as an aid to the reader. Arrowheads along the X-axis (the anteroposterior extent of AI) indicate the approximate orientation of the isofrequency contours. The Y-axis represents the dorsoventral extent of the map, which in this case measured 2.2 mm. The numbers in the box below the pseudo three-dimensional

#87-420

Fig. 6. The spatial distribution of preferred speed as determined by the non-quantal measurement. The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of preferred speed. Two dimensions of the projection represent the anteroposterior (a,p) and dorsoventral (d,v) extent of the cortical surface. The *arrowheads* along the anteroposterior dimension indicate the approximate orientation of the isofrequency contours. The elevation of the surface of the plot corresponds to the preferred speed across the cortical surface. A smoothing factor of 0.9 was applied to the data, resulting in a small amount of additional smoothing (see Schreiner and Mendelson 1990). The *numbers* along the elevational axis refer to the maximum and minimum preferred speed values for this case. The *low number*, and consequently the low elevations, represents a preference for slow FM sweeps and the *high number* and coincident peak indicate a preference for fast FM sweeps. The *dotted line* indicates a preference for medium FM speeds (indicated by the *number* 2) and has been included to aid in interpreting the maps. As illustrated, cells that preferred fast speeds were located dorsally and were followed by a region where cells preferred medium speeds. In the ventral portion, cells responded best to slow FM sweeps. In the *lower portion*, the *numbers in the box* represent the actual obtained preferred speed response of FM sweep for a given location, as determined by the non-quantal measurement

Fig. 7. A comparison of the spatial distribution of preferred FM speed determined with the non-quantal (A) and the discrete measurement (B). Both figures are pseudo three-dimensional plots of preferred speed calculated in these two ways (see Materials and methods). Little difference between the two methods was apparent in the spatial distribution of preferred speed. Conventions are identical to those in Fig. 6

projection represent the preferred speed of FM sweep for a given location as determined by the nondiscrete measurement technique described in Materials and methods. In the dorsal aspect of this map, recording locations were marked by responses showing a preference for fast FM sweeps, while in the more ventral region, slower FM sweeps were preferred. Responses to medium-speed FM sweeps were encountered in an intermediate region between fast and slow speed response subregions. This spatial distribution for preferred FM speed was similar across the range of frequencies examined.

Figure 7A shows the distribution of preferred speed for case 87-518 calculated by the nondiscrete classification method, and Fig. 7B illustrates the distribution determined by the discrete measurement technique of this response parameter (see Materials and methods). The spatial distributions derived from both of these methods were similar to each other across all cases examined. The spatial distribution of preferred FM speed for this case differed somewhat from other cases. Sites located in the

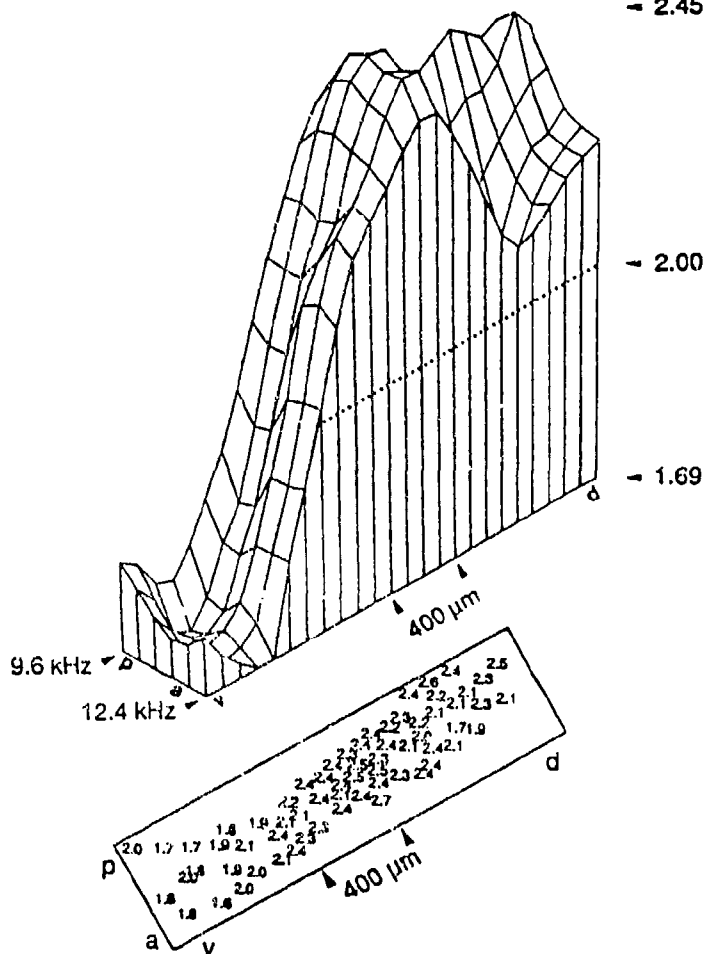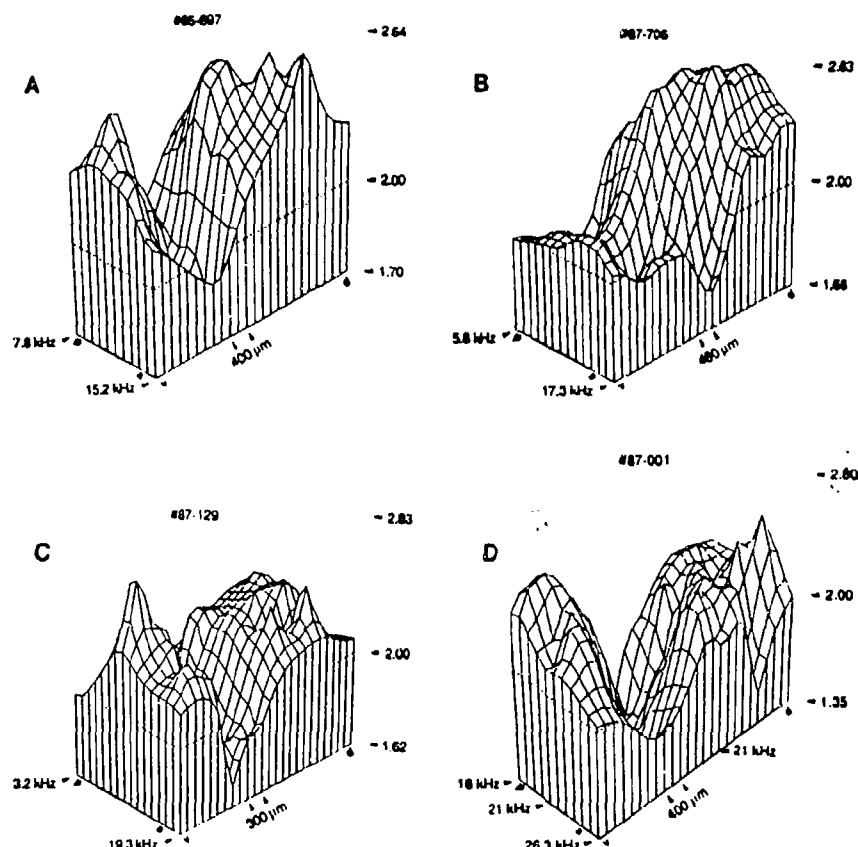Fig. 8A–D. Additional examples of the spatial distribution of preferred speed as determined by the non- quantal measure. Pseudo three-dimensional representation of preferred speed for case 86-697 (A), case 87-706 (B), case 87-129 (C), and for case 87-001 (D). Note that the orientation of the mapped area for case 87-001 (D) was tilted 24° relative to the isofrequency orientation, resulting in an approximate alignment of the 18.5-kHz contour with the ventroposterior to dorsoanterior axis of the map. In all four cases a systematic distribution of preferred FM speed was observed. Conventions are identical to those in Fig. 6

more dorsal and ventral regions were distinguished by preferences for medium as opposed to fast FM sweep speeds, while for the central region interposed between these extreme dorsal and ventral subregions, the more typical dorsoventral spatial pattern of fast-medium-slow-medium-fast was observed. However, in concurrence with the other cases, the preferred FM speed responses for the map illustrated in Fig. 7 appeared to be largely independent of CF (7.0 kHz < CF < 11.0 kHz).

As there was little difference between the spatial distribution of preferred speed responses as determined by the nondiscrete and discrete measurement techniques, only the former measurements will be illustrated in the remaining cases. To demonstrate the consistency of the findings, as well as evidence for interindividual variations, Fig. 8 shows four additional cases of the distribution of preferred FM speed responses in AI..All of the maps contained a circumscribed, elongated depression corresponding to slow FM sweep speed preferences flanked by a series of gradual transitions to regions containing preferences for medium or fast speeds. Generally, the slow-speed preferences were located between the lower and middle third of the mapped dorsoventral dimension and were oriented orthogonally to the isofrequency contours. Figure 8A (86-697) provides an example of the spatial distribution of preferred FM speed for sites with CFs between 7.8 and 15.2 kHz. In the most dorsal aspect of this map, preferences for fast and/or medium FM rates were observed. Adjacent to this region, indicated by a peak in the map, cortical locations were marked by preferences for fast FM sweeps. A second maximum (i.e., sub-

region of fast FM sweep selective units) was located in the more ventral aspect of the map. In the middle of the map (represented by the through), responses exhibited a preference for slow speeds. Interpolated between the fast and slow subregions were recording sites consisting of responses to medium FM sweeps. In the dorsal and central regions, preferences for fast or medium FM sweeps appeared to be independent of CF. In the ventral region, however, the selectivity for fast FM sweeps was expressed more strongly for lower CFs (e.g., around 7–8 kHz) while locations with higher CFs (e.g., 15.0 kHz) appeared to prefer medium sweep rates.

Figure 8B (87-706) shows another example in which recording from several sites (5.8 kHz < CF < 17.3 kHz) in the dorsal and middle regions of AI yielded a distribution of responses, preferences for fast and medium FM sweeps, respectively. In contrast to other maps, the majority of sites in the ventral aspect of this map were characterized by preferences for medium rates of FM sweeps, despite the fact that the dorsoventral extent explored in this case (5 mm) was comparable with most other cases.

The preferred FM speed distribution shown in Fig. 8C (87-129; 3.2 kHz < CF < 19.3 kHz) is similar to the distribution illustrated in Fig. 8A in that only the ventral representation of FM speed responses appeared to vary with CF. Those recording sites where neural response preferences for fast speeds were encountered were also marked by CFs between 7.0 and 9.0 kHz, while lower or higher CFs appeared to be associated with preferences for medium-speeds.

In Fig. 8D (87-001) the orientation of the mapped area

was tilted 24° relative to the isofrequency orientation (see Fig. 4 in Schreiner and Mendelson 1990), resulting in an approximate alignment of the 18.5-kHz contour with the ventroposterior to dorsoanterior axis of the map. Considering only the most complete isofrequency contours (18.5–22.0 kHz) it becomes apparent that the preferred FM speed followed the same general pattern of spatial distribution (i.e., fast-medium-slow-medium-fast) seen in most other cases.

In summary, preferences for FM sweeps of different speeds seemed to be systematically distributed along the dorsoventral axis of the high-frequency region of AI. In the most dorsal aspects of AI, the majority of recording sites were distinguished by responses showing a preference for fast and/or medium FM sweeps. In this region, approximately equal responses to both fast and medium FM sweeps were also encountered. If one considers the progression of FM speed responses encountered in electrode penetrations running along the dorsoventral axis of AI, responses alternated in the following way: first, responses encountered consisted of preferences for fast speeds, then medium-speeds, and then slow speeds, followed by responses that preferred medium speeds again, until in the most ventral region there was a tendency for multiunits to prefer fast FM sweeps. In the more ventral aspect of AI (and possibly including the dorsal region of AII), more variability between maps was observed. In the majority of cases, this spatial pattern of preferred speed distribution appeared to be oriented orthogonally to the isofrequency domain of AI and could be considered to be largely independent of CF.

Representation of direction selectivity

The distribution of preferred responses to each direction of FM sweep for individual cases and for all cases combined is shown in Table 2. In general, at just over half of the cortical locations examined (55.4%), no preference for the direction of FM sweep was observed (using a criterion of ±0.15; see Materials and methods). In one case (87-129), almost 75% of the recording sites contained responses classified as non-direction-selective. Of the total number of recording sites for all cases where direction selective responses were found, almost twice as many preferred downward- to upward-directed FM sweeps. For two out of six cases (87-706 and 87-129), the reverse trend was observed, consisting of a slightly higher percentage of responses preferring upward- as opposed to downward-directed FM sweeps.

Figure 9 shows a scatter plot of direction selectivity as a function of CF for two cases. The two dotted lines, corresponding to ±0.15 (see Materials and methods), represent the chosen classification boundaries for direction-selective responses. Thus, points falling between these lines were classified as non-direction-selective while those falling outside were classified as direction selective (negative numbers represent downward-directed responses; positive numbers represent upward-directed responses). While the range of CFs tested for case 87-518
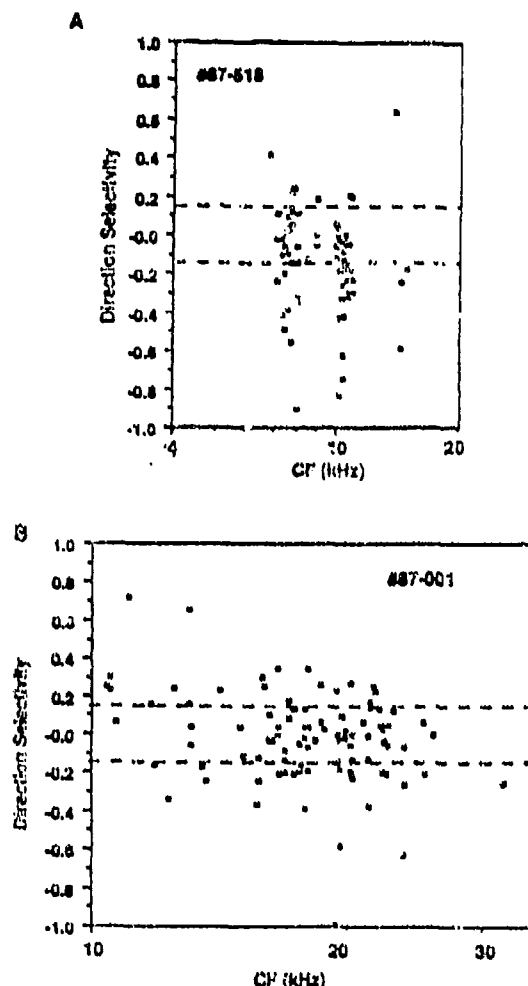
Fig. 9A,B. Scatter plots of direction selectivity as a function of characteristic frequency (CF). All of the direction-selective index values are plotted as a function of CF for case 87-518 (A) and 87-001 (B). Positive values refer to a preference for upward-directed FM sweeps, while the negative values refer to a preference for downward-directed FM sweeps. The dotted lines indicate the boundaries (±0.15) for direction-selectivity classification. Symbols between the dotted lines represent unit clusters that were non-direction-selective. In general, the degree of direction selectivity appeared to be independent of CF

Table 2. Distribution of preferred direction of frequency-modulated sweeps

Cat	Number of sweeps	Upward (%)	Downward (%)	No preference (%)
86-697	88	9.1	42.0	48.0
87-420	63	15.9	27.0	57.1
87-518	82	8.5	50.0	42.5
87-706	86	23.3	17.4	59.3
87-129	81	14.8	11.1	74.1
87-001	96	20.8	26.0	53.1
Total	496	15.5	29.0	55.4
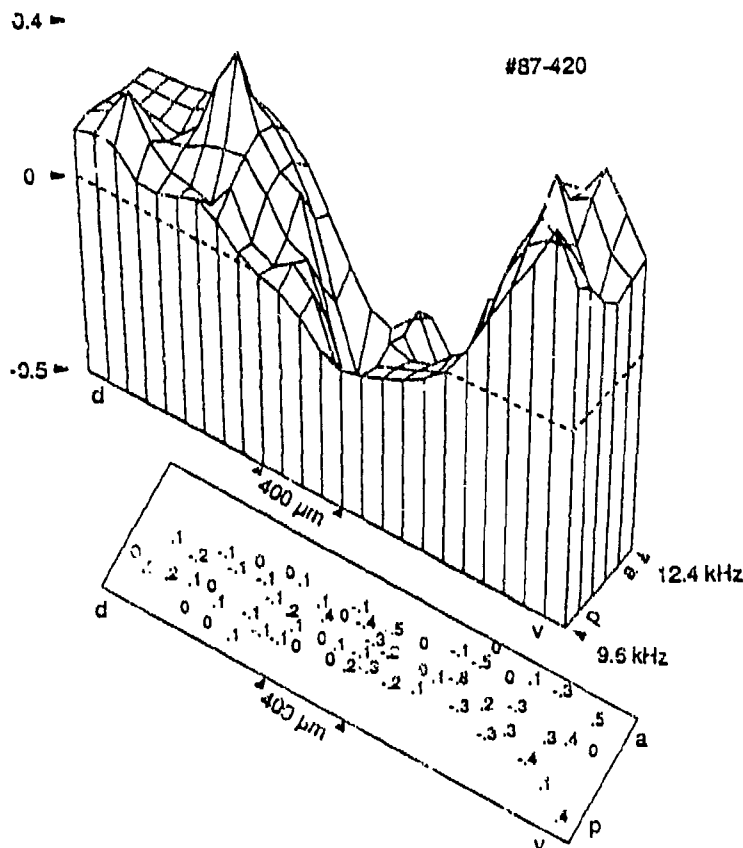
Fig. 10. The spatial distribution of direction-selective responses. As with the pseudo three-dimensional plots depicting preferred speed, in the *upper portion* of the figure two dimensions of the projection represent the anteroposterior (*a,p*) and dorsoventral (*d,v*) extent of the cortical surface. The *arrowheads* along the anteroposterior dimension indicate the approximate orientation of the isofrequency contours. The elevation of the surface of the plot corresponds to the direction selectivity across the cortical surface. An additional smoothing factor of 0.9 was applied to the data. The *numbers* along the elevational axis refer to the direction-selective index values for this case. The *negative number* and corresponding low elevations represent a preference for downward-directed FM sweeps. The *positive number* and corresponding peaks represent a preference for upward-directed FM sweeps. The *dotted line* represents a direction-selective index value of 0. In the *lower portion*, the *numbers in the box* represent the actual direction-selective index values obtained for each location. For this case, cells that responded best to upward-directed FM sweeps were located in the dorsal and the ventral regions, while downward-directed selective cells were located in the central region

(Fig. 9A) was much narrower than for case 87-001 (Fig. 9B), both scatter plots suggest that direction selectivity was independent of CF.

In general, it was found that preferences for one FM sweep direction were spatially segregated from responses for the opposite direction. Figure 10 (case 87-420) shows an example of the spatial distribution of direction selectivity across AI. As in the pseudo three-dimensional plots for preferred FM speed (Figs. 6–8), the X- and Y-axes represent the anteroposterior and dorsoventral extent of AI, respectively. In addition, surface elevation (Z-axis) represents the direction-selective index value obtained at each cortical location examined within this area. The dotted line represents a direction-selective index value of 0. Below the pseudo three-dimensional plot is a box representing the cortical locations sampled. The numbers in the box indicate the direction-selective index value computed for responses observed at each location (negative values represent a preference for downward-directed FM sweeps and positive values represent a preference for upward-directed FM sweeps). In Fig. 10, neurons located in the dorsal region showed little or no direction selectivity. Responses recorded around the 10 kHz region displayed a weak preference for upward-directed FM sweeps. In the mid-region, a higher incidence of selectivity for downward-directed sweeps was observed, while in the most ventral aspect of this map, most units exhibited a preference for upward-directed FM sweeps.

Figure 11 shows three additional examples of the spatial distribution of direction selective responses across AI. Figure 11A (87-518) shows a map in which preferences for upward-directed FM sweeps were confined exclusively to the dorsal region of AI. In the ventral region, most responses were non-direction-selective. In the middle region, preferences for downward-directed FM sweeps were prevalent. For this example, direction- and non-direction-selective responses appeared to be independent of CF.

The map illustrated in Fig. 11B (86-697) is unique in that there were three distinct regions of downward-direction selective FM responses and three regions distinguished by a concentration of upward direction-selective responses. Further, the range of direction-selective index values (−0.82 to +1.0) was slightly greater than for other cases. In the dorsal and mid-regions of the surface map, direction selectivity appeared to be dependent upon CF. Upward-directed responses were found at higher CFs (around 13.0–15.0 kHz) in the dorsal region and at lower frequencies (8.0–10.0 kHz) in the middle region.

In the dorsal region of case 87-001 (Fig. 11C) cells with lower CFs responded best to upward-directed FM sweeps, while cells with higher CFs responded best to downward-directed FM sweeps. In the mid-region of the map, responses were strongly selective for downward-directed FM sweeps and were largely independent of CF In the ventral aspect of the map, a preference for down-

A #87-518
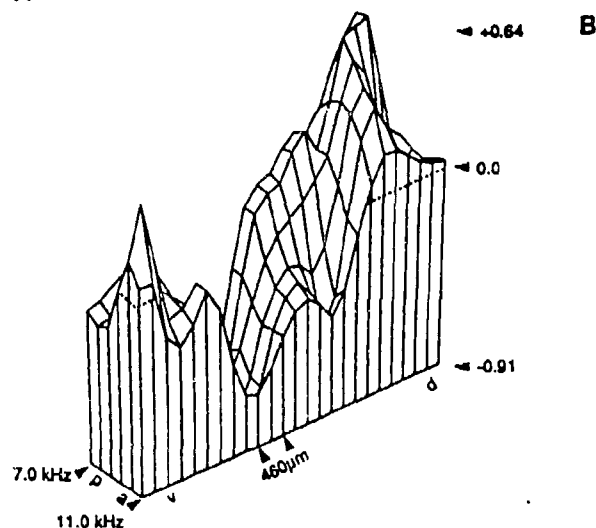
B #86-697
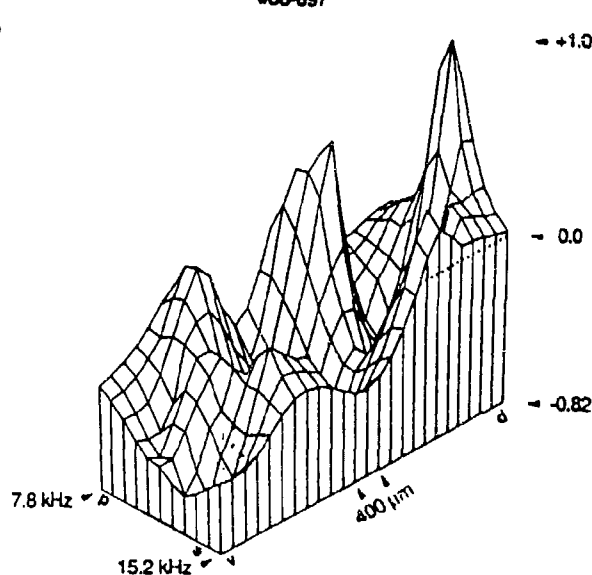
C #87-001
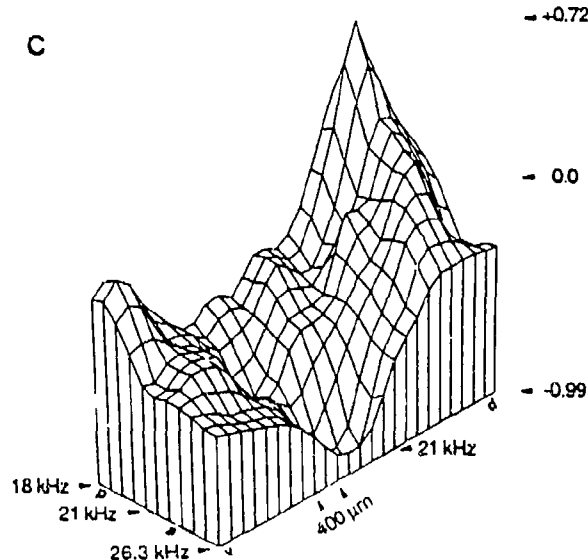
Fig. 11A–D. Additional examples of the spatial distribution of direction selectivity. A Pseudo three- dimensional representation of direction selectivity for case 87-518. B Spatial distribution of direction selectivity for case 86-697 represented as a pseudo three-dimensional projection. C Pseudo three-dimensional representation of direction selectivity for case 87-001. Note that the orientation of the mapped area for this case was tilted 24° relative to the isofrequency orientation. All three cases showed a nonrandom spatial distribution of direction-selective FM sweep responses. Conventions are identical to those in Fig. 10

ward-directed FM sweeps was also apparent, particularly for units with CFs between 15.0 and 26.3 kHz, although the degree of direction selectivity was reduced relative to the middle region. Low-frequency sensitive responses encountered in the ventral posterior region of this map were not selective for FM sweep direction.

A comparison of all the cases shown in Figs. 10 and 11 suggests that the spatial organization of direction-selective FM responses may obey similar rules. However, the variability in the global representation of these response properties appears to be greater than was observed for FM sweep speed.

In summary, examination of directional preferences for FM sweeps over a wide area of cortex showed evidence of a systematic distribution along the dorsoventral axis of the high-frequency region of AI. In the dorsal region of AI, most direction-selective response prefer-

ences occurred for upward-directed FM sweeps, while in the middle region the majority of responses showed strong preferences for downward-directed FM sweeps. These directional preferences generally appeared to be independent of CF, although some non-uniformity in the distribution along the frequency gradient was evident. Toward the ventral aspect of AI, the preference for downward-directed sweeps gradually declined leading to either no direction selectivity or to a preference for upward-directed FM sweeps. In some ventral locations, the directional preference seemed to be dependent upon CF.

Direction selectivity as a function of rate of FM sweep

In the previous section, direction selectivity was associated with maximal responses (see Materials and methods)
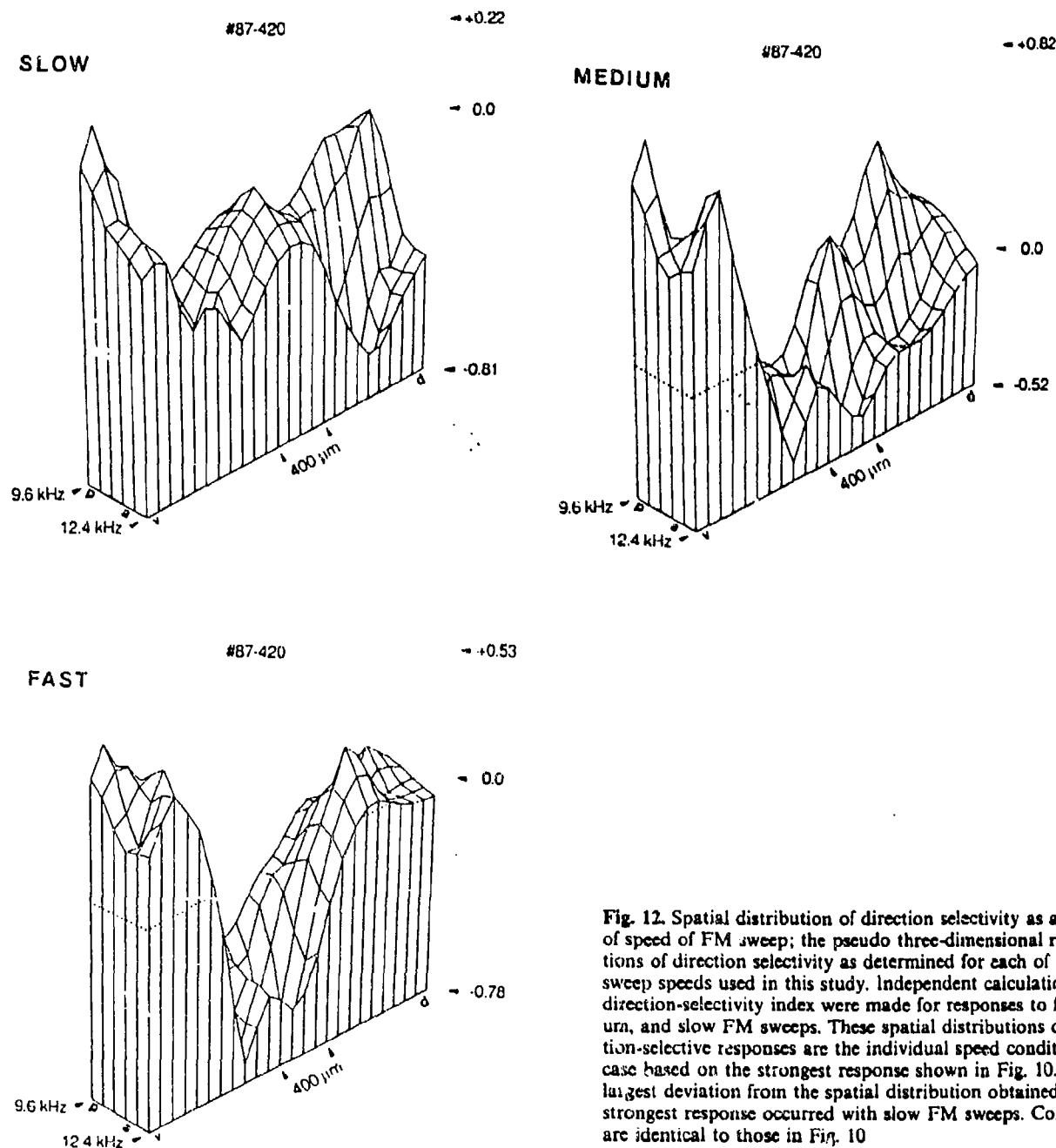
Fig. 12. Spatial distribution of direction selectivity as a function of speed of FM sweep; the pseudo three-dimensional representations of direction selectivity as determined for each of the three sweep speeds used in this study. Independent calculations of the direction-selectivity index were made for responses to fast, medium, and slow FM sweeps. These spatial distributions of direction-selective responses are the individual speed conditions of the case based on the strongest response shown in Fig. 10. The largest deviation from the spatial distribution obtained with the strongest response occurred with slow FM sweeps. Conventions are identical to those in Fig. 10

and thus was considered without specific reference to the possible effects FM sweep speed might play in the generation of direction-selective responses. In this section, the distribution of direction-selective responses is re-examined using three different speeds of FM sweep. Thus, the contribution of sweep speed to direction selectivity of FM sweep could be estimated by calculating independent direction-selective index values for each set of responses to fast, medium, and slow FM sweeps for each case.

The spatial distribution of direction-selective responses illustrated in Fig. 12 (87-420) was obtained using the same speed conditions that were used to generate the results shown in Fig. 10. At medium and fast speeds, the more common bimodal distribution of upward direction-selective responses was apparent. For the medium-speed, this preference for upward-directed FM sweeps was more frequently observed in the dorsal and especially in the ventral region when compared with the distribution of direction selectivity based on the strongest response ($+0.82$ in the medium-speed plot vs $+0.49$ in the strongest-response plot). The cortical distribution of direction-selective responses assessed with fast FM speeds also differed from the distribution based on the strongest-response criterion in the dorsal region, where direction-selective responses were less frequently encountered. The preferences for downward-directed FM sweeps concen-

SLOW

#87-001

— +0.63

— 0.0

— -0.93

18 kHz —
21 kHz —
26.3 kHz —
21 kHz
400 μm

MEDIUM

#87-001   — +0.87

— 0.0

— -0.99

18 kHz —
21 kHz —
26.3 kHz —
21 kHz
400 μm

FAST

#87-001

— +0.59

— 0.0

— -0.63

18 kHz —
21 kHz —
26.3 kHz —
21 kHz
400 μm
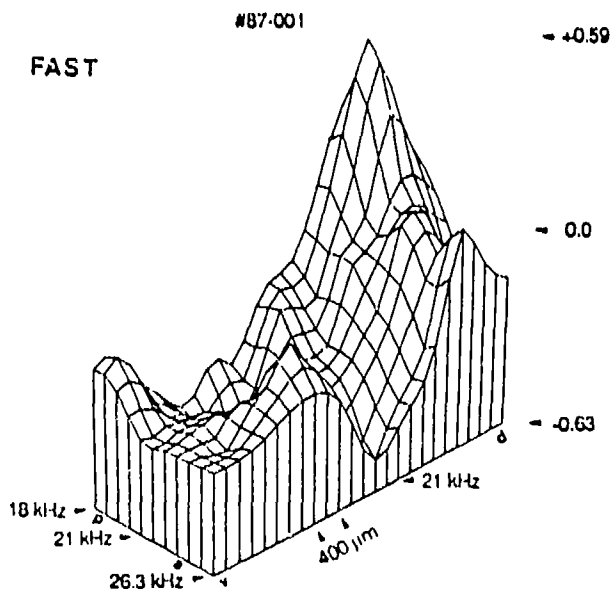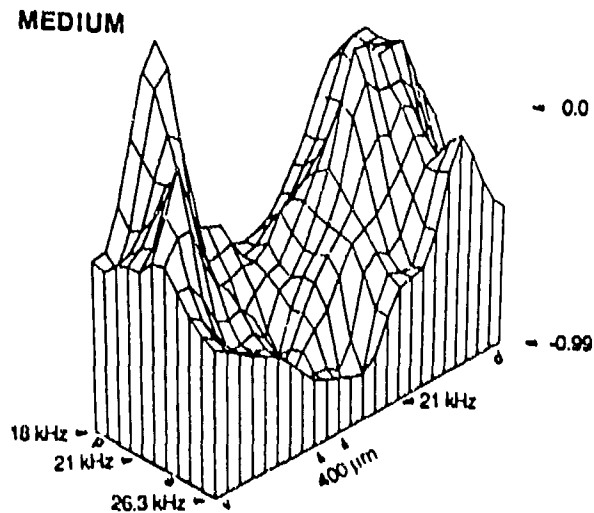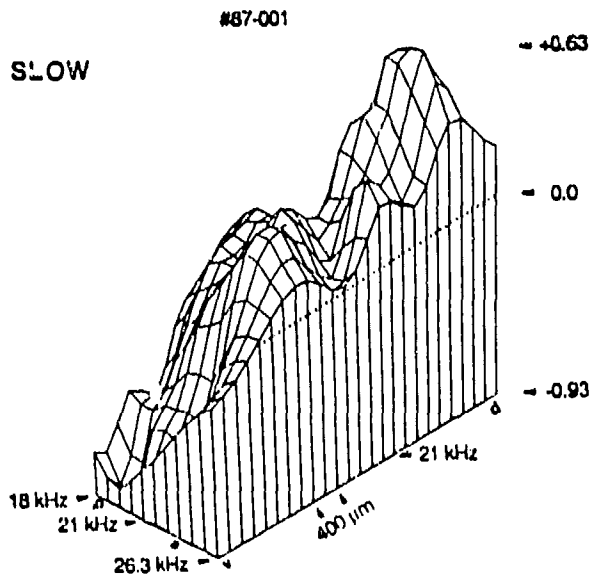
Fig. 13. A second example of the spatial distribution of direction selectivity as a function of speed of FM sweep. The three pseudo three-dimensional representations of direction selectivity based on the individual speeds are for the strongest response case shown in Fig. 11C. Note that the orientation of the mapped area for this case was tilted 24° relative to the isofrequency orientation, resulting in an approximate alignment of the 18.5-kHz contour with the ventroposterior to dorsoanterior axis of the map. The spatial distribution of direction selectivity obtained with slow FM sweeps differed the most from that obtained with the strongest response. For conventions see Fig. 10

trated in the mid-region of AI, as determined using both medium and fast FM sweep rates, was similar to the distribution based on the strongest response. The distribution of direction-selective responses for slow FM sweeps differed from the strongest response distribution in three respects. First, relatively few multiunit responses at any site showed a preference for upward-directed sweeps. Second, responses at sites in the dorsal region where responses displayed CFs ranging from 11.0 to 12.0 kHz also exhibited strong preferences for downward-directed FM sweeps. Third, the magnitude of direction selectivity was reduced at some of the same locations where responses showed marked preferences for fast and medium speeds.

Figure 13 (87-001) shows another example of the dis-

tribution of direction-selective responses as a function of the individual FM speed conditions (see Fig. 11C). Of the three speeds examined, the distribution of direction-selective responses tested with fast FM sweeps more closely resembled the distribution of direction selectivity based on the strongest response. The plots for medium and slow FM sweeps differed most in the ventral region from the direction selectivity distribution of the strongest response. As with the case shown in Fig. 12, the distribution of direction-selective responses tested with slow FM sweeps differed most from the other two FM speed distributions as well as from that derived from the strongest response.

In summary, the spatial distribution of direction-selective responses to FM sweeps depended on the sweep

A
#87-518
DS < -0.16   PS > 2.2

B
#87-001
DS < -.02   PS > 2.0
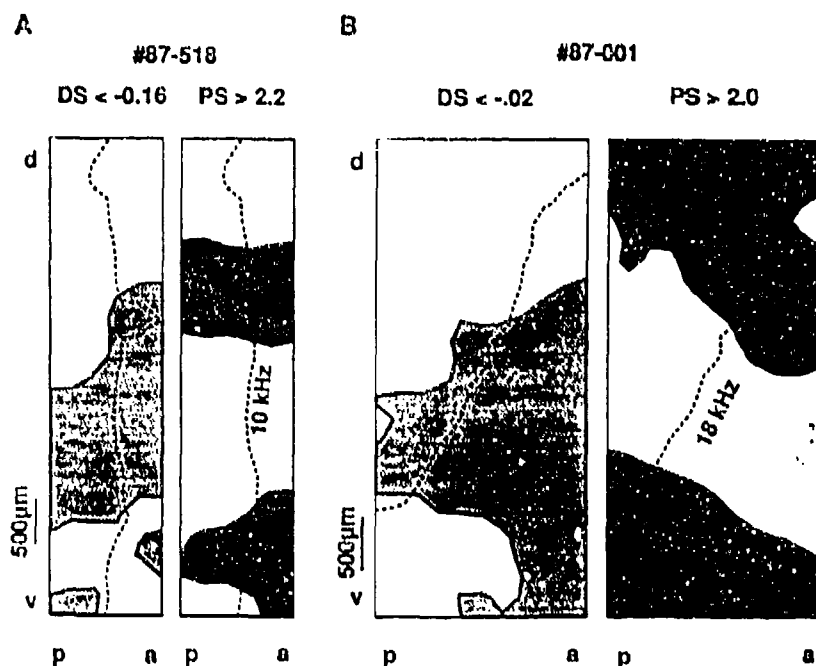
Fig. 14A,B. Comparison of direction selectivity (*DS*) and preferred speed (*PS*). The contour plots on the *left* represent the spatial distribution of direction selectivity. The *hatching* depicts those cortical regions where units responded best to downward-directed FM sweeps (DS index <-0.16 and DS index <-0.2, respectively). The contour plots on the *right* represent the spatial distribution of the preferred FM speed. The *hatching* depicts those cortical areas where locations preferred fast FM sweep speeds (PS index >2.2 or 2.0, respectively). The two dimensions of the plot represent the anteroposterior (*a,p*) and dorsoventral (*d,v*) extent of the cortical surface. A Direction-selective and preferred-speed contour plots for case 87-518. The *dashed line* depicts an isofrequency contour of 10.0 kHz. B Direction-selective and preferred-speed contour plots for case 87-001. The isofrequency contour for 18.0 kHz is represented by the *dashed line*. The orientation of the mapped area for this case was tilted 24° relative to the isofrequency orientation. Note that the nearly complementary distribution of locations with preference for downward-directed sweeps and fast sweep rates was observed in both cases

speed used to elicit the response and differed somewhat from the distribution based on the strongest response. Direction-selective response distributions determined with fast and medium rates of FM sweeps showed closer agreement with those based on the strongest response criterion, especially in the dorsal and middle regions of AI. However, the greatest difference between the spatial distributions of direction-selective responses assessed using these criteria was observed in the case of slow FM sweeps. Particularly notable in the slow FM speed map was a reduction in the degree of direction selectivity for the more centrally located responses compared with the strong, downward direction-selective responses that occupy this region of the map when faster FM speeds are used.

Correlation of direction selectivity and preferred speed

An attempt was made to correlate the spatial distribution of preferred speed with that of direction selectivity. The contour plots in Fig. 14 illustrate the relationship between the spatial distributions of direction-selective and preferred-speed responses for two cases. The dashed line represents the orientation of an isofrequency contour running through the center of each map. In the two cases illustrated, 87-518 (Fig. 14A) and 87-001 (Fig. 14B), the contour plots on the left of each panel show the spatial distribution of direction-selective responses while the plots on the right show the spatial distribution of preferred FM speed responses. The shaded areas in the contour plots for direction selectivity (DS) demarcate those regions where multiunit responses reflect a preference for downward-directed FM sweeps (DS <-0.16 and DS <-0.2, respectively) while the shaded areas in the pre-

ferred FM speed (PS) plots illustrate regions where the best response was to fast FM sweeps (PS >2.2 and PS >2.0, respectively). As can be seen in Figs. 14A and B, there was little overlap between the areas containing downward- direction-selective and fast speed responses to FM sweeps for the two cases. This observation supports the notion of a covariance of direction selectivity and preferred speed: recording sites in which multiunit responses prefer fast FM sweeps coincide with locations of multiunit responses preferring upward-directed FM sweeps and vice versa. However, significant deviation from this pattern may be seen in the upper third and in the ventral extremes of the mapped areas. This suggests that the covariance between direction and speed selectivity may not be homogeneously distributed across the entire cortical area examined by these studies. This is also supported by the results of a point-by-point correlational analysis of direction and speed selectivity. Table 3 shows that no statistically significant correlation could be found between direction and speed selectivity across all points in each map. Regional correlations for preferred speed and direction were carried out independently in the dor-

Table 3. Correlation of preferred speed and preferred direction of frequency-modulated sweep

Cat	Number of sweeps	Correlation (r)	P
86-697	88	0.07	0.51
87-420	63	0.04	0.78
87-518	81	0.08	0.49
87-706	83	0.02	0.85
87-129	81	0.20	0.07
87-001	96	0.12	0.25

sal (i.e., the portion of AI dorsal to the most sharply tuned multiple unit responses; see Schreiner and Mendelson 1990) and ventral (i.e., locations ventral to the most sharply tuned region of AI) subregions of each map also yielded no statistically significant correlations between these response properties.

Correlation of direction selectivity and preferred speed with excitatory bandwidth and broadband transient response

In general, there was a tendency for multiunit responses recorded in the more dorsal and ventral regions of AI to respond best to fast or medium FM sweeps changing in an upward direction, while multiunits examined in the middle region responded best to downward-directed and slower FM sweeps. In a previous paper, Schreiner and Mendelson (1990) showed that in the dorsal and ventral regions of AI, cells had low Q_{10dB} and Q_{40dB} values and responded best to broadband transient stimuli. Cells located in the middle region were, in contrast, narrowly tuned and responded poorly to broadband stimuli. Thus, in an attempt to reveal the neural mechanisms underlying FM sweep responses, the relationships between excitatory bandwidth and broadband transient response with preferred FM sweep speed and direction were examined. To this end, statistical correlations between these response parameters were analyzed separately for cells located in the dorsal region (AId) and the ventral region (AIv) of the auditory cortex. In addition, statistical correlations were also examined for responses sampled over

Table 4. Correlation of preferred frequency-modulated speed with integrated excitatory bandwidth and response strength to broadband transient signals

Case	AI		AId [a]		AIv [b]	
	r	P	r	P	r	P
Q_{10dB}						
87-518	−0.26	0.03	−0.38	0.01		
87-706	−0.20	0.05				
87-001	−0.24	0.02	−0.31	0.02		
Q_{40dB}						
87-518	−0.56	0.0001	−0.50	0.001	−0.58	0.0002
87-001	−0.29	0.005	−0.29	0.02		
BTS						
87-518	0.47	0.0001	0.58	0.0001	0.40	0.01
87-706	0.22	0.05				
87-001	0.45	0.0001	0.44	0.0005	0.44	0.008

Linear regression analysis: r, correlation coefficient; P, level of significance (F-test)

[a] AId represents recording locations in the portion of AI dorsal to the most sharply tuned multiple unit responses

[b] AIv represents locations ventral to the most sharply tuned region of AI (Schreiner and Mendelson 1990)

BTS, broadband transient signals; AI, primary auditory cortex; Q, excitatory bandwidth

Table 5. Correlation of direction selectivity for frequency-modulated sweep and integrated excitatory bandwidth

Case	AI		AId		AIv	
	r	P	r	P	r	P
Q_{10dB}						
87-518	−0.25	0.025	−0.33	0.03		
87-706	−0.30	0.005	−0.36	0.01		
87-001	−0.33	0.001	−0.32	0.01	−0.38	0.002
Q_{40dB}						
87-518	−0.49	0.0001	−0.36	0.02	−0.57	0.0003
87-001	−0.29	0.004	−0.32	0.01		

Linear regression analysis: r, correlation coefficient; P, level of significance (F-test). For further explanation see Table 4

the entire cortical region (AI) in the three cases for which complete Q_{10dB} and Q_{40dB} data were available.

The correlation of integrated excitatory bandwidth and broadband transient response with preferred FM speed is shown in Table 4. Only statistically significant correlations have been indicated. As can be seen, there was a strong correlation between preferred FM speed and excitatory bandwidth across the entire region of cortex examined in these experiments. This correlation indicates that units with higher Q_{10dB} and Q_{40dB} (i.e., units that were narrowly tuned) respond best to slower FM sweeps and vice versa. When examined as a function of cortical subregion, there was a stronger correlation between these response parameters and preferred FM speed in the dorsal than in the ventral subregion. For broadband stimulus response and preferred FM sweep speed, a strong correlation was also observed in both halves of the cortical regions studied. Thus, units that responded best to faster speeds of FM sweeps also seemed to respond well to a broadband transient stimulus and vice versa. In all three cases, there was a significant correlation between FM direction selectivity and integrated excitatory bandwidth over the entire extent of the mapped cortical region (Table 5). In particular, significant correlations were observed for multiunit responses obtained in the dorsal rather than the ventral region of the cortex. Unlike preferred speed of FM sweep, there was no significant correlation between direction selectivity and responses to broadband transient signals.

Discussion

The primary purpose of the present study was to determine the systematic spatial distribution of neural responses in the cat AI to the direction and speed of FM sweeps. For these two FM sweep parameters, a nonrandom distribution along the isofrequency dimension was observed. Extensive mapping of area AI revealed that response preferences to fast and/or medium FM speeds were usually encountered dorsally, slow FM speed-selective responses were found centrally, and fast speed-selective responses were located ventrally within AI. Medium-

speed selective responses were juxtaposed between the fast and slow speed selective regions, supporting the conclusion that there is an orderly, functional transition in speed tuning along the isofrequency domain of AI. For most direction-selective responses, preferences for upward-directed FM sweeps were found in dorsal and, to a lesser degree, in ventral subregions of AI as well, while preferences for downward-directed FM sweeps were found in the central region of AI. Although the size and extent of speed- and direction-selective response regions varied to some extent across cases, the overall distribution of these response parameters within AI suggests that in addition to tonotopicity, the functional organization of AI may also be defined by an orthogonal dimension of FM sweep responses within the isofrequency domain.

Before discussing these results further, two preliminary issues should be addressed. First, since multiunit recording samples the activity of several neurons within a small volume, it cannot be assumed that all neurons involved have identical response properties (Schreiner and Mendelson 1990). Furthermore, the contribution of thalamocortical afferents to local cortical responses, although probably minimal, cannot be completely excluded (Schreiner and Cynader 1984). Therefore, cortical multiunit responses should be interpreted as a local extracellular manifestation of integrated cortical processes operating at a given location rather than as a faithful reflection of underlying single-unit responses (Schreiner and Mendelson 1990). Accordingly, both direction-selective and speed-selective multiunit responses recorded at a given cortical site reflect the dominant pattern of neural activity resulting from local cortical processing at that location. The systematic variation in FM-selective multiunit responses observed along the dorsoventral extent of AI is more easily explained as the consequence of intrinsically distributed cortical processes rather than as the result of random variation in multiunit responses. Therefore, because FM-selective responses varied systematically as a function of cortical location, it is more likely that these variations reflect the global operation of extensive cortical networks, rather than location-independent, random variations in multiunit activity (Schreiner and Mendelson 1990).

The second issue is concerned with the possible systematic influences of the stimulus delivery system. Irregularities in the stimulus delivery system may create potential contributions to the recorded responses such as, for example, AM components in the delivered signal. Given the broad FM sweeps used in the present study (from 0.25 to 64.0 kHz) and the known irregularities in the transfer function of the sound delivery system (see Materials and methods), it is likely that the FM sweeps used in these experiments also contained an AM component, especially at higher frequencies. However, as pointed out by Suga (1968), the effect of the potentially confounding amplitude component on the neural responses encountered is likely to be minimal, since the possible AM contributions would be similar for most stimuli (eg., increasing AM component with a downward-directed FM sweep and decreasing AM component with an upward-directed FM sweep), while neural responses to FM stim-

uli, observed in the present study, systematically differed from one another.

Furthermore, if our FM data were greatly confounded by an AM component, then, based on the results of Phillips and Hall (1987), one would predict that the majority of cortical cells would respond most vigorously to downward-directed FM sweeps. This prediction is based on the premise that the high-frequency roll-off in the transfer function of most sound delivery systems implies that downward-directed FM sweeps will occasion a stimulus that increases progressively in amplitude as it decreases in frequency. Phillips and Hall (1987) showed that, with the exception of cells with off responses, most cells (87%) responded to increases in stimulus intensity. However, 55% of the unit clusters sampled in the present study did not respond selectively to either direction of FM sweep. In addition, of the direction-selective multiunit responses, some preferred upward-directed while others preferred downward-directed FM sweeps. Thus, these considerations strongly suggest that nonspectral contributions to the FM-selective responses reported here were minimal and, therefore, do not seriously compromise the interpretation or the generality of these findings.

Cortical depth and FM sweep selectivity

Cells recorded along a single electrode penetration (oriented perpendicular to the cortical surface) displayed similar preferences for the speed and direction of FM sweeps. This was clearest at depths between 650 and 2000 μm below the cortical surface. Despite the relatively small number of long penetrations investigated in the present study, the similarity of responses recorded at diverse depths suggests that FM speed- and direction-selective responses are column-dependent properties. This is consistent with the single-unit study of Mendelson and Cynader (1985) in which they reported that cells within an electrode penetration perpendicular to the cortical surface exhibited similar FM sweep speed and direction-selective responses. Furthermore, other response properties such as CF, binaural interaction, time-delay sensitivity, and threshold have also been found to exhibit similar columnar-dependent properties (Suga 1965a; Brugge and Merzenich 1973; Imig and Adrian 1977).

Stimulus level and FM sweep responses

The effect of stimulus level on preferred speed and direction selectivity was statistically significant. This is perhaps not surprising given that excitatory bandwidth tuning is also known to be affected by stimulus level (Schreiner and Mendelson 1990). In addition, as we have demonstrated here, integrated excitatory bandwidth is significantly correlated with preferred speed and direction selectivity (see Tables 4, 5). Although significant, the slopes of the linear regression lines fitted to both preferred-speed and direction-selectivity data were relatively shallow. This suggests that the trend toward faster speed selectivity and downward direction selectivity for FM

sweeps with increasing stimulus level was minimal. It should be noted that Moller (1974) also observed a higher occurrence of downward-directed cells in the rat cochlear nucleus as stimulus intensity was increased.

Preferred speed

The cortical distribution of preferred speed responses exhibited steep gradients along the dorsoventral axis of AI. In many cases, the distribution of preferred speed responses was largely independent of CF, especially in the dorsal and central subregions of AI. In cases where preferred speed varied with CF (see Fig. 8), variability was greatest in the ventral aspect of AI. A possible explanation for this variability is that some ventral recording sites may actually have been located in the dorsal aspect of the adjacent auditory field (AII). Neurons in AII are characterized by broad tuning curves (i.e., with small Q_{10dB} values; see Schreiner and Cynader 1984; Schreiner and Mendelson 1990). However, the transition from AI to AII as measured by the progression of Q_{10dB} values is extremely gradual such that the distinction between these neighbouring cortical regions may be nondetectable (Schreiner and Cynader 1984). Whatever cortical field they occupied, multiunit responses in the ventral region appeared to be functionally distinct from responses encountered at more dorsal cortical locations in AI. Additional support for this functional distinction is derived from the stronger correlations observed between the integrated excitatory bandwidth and preferred FM speed in the dorsal as compared to the ventral regions of AI.

Although the three FM speeds used in the present study were not sufficient to uniquely determine the preferred FM speed for each location, they did allow a global assessment of the spatial distribution of speed preferences across the auditory cortex. Previous studies (Mendelson and Grasse 1992) of FM speed tuning, together with single-unit control experiments in this study (see Fig. 4), indicate that the FM speeds selected in these experiments were sufficiently spaced to safely avoid undersampling speed-selective responses. However, it cannot be completely ruled out that some of variability in the data could be due to cells or cell clusters contributing to the speed-tuning responses which themselves prefer speeds that fall either completely outside, or somewhere in between the actual speed conditions used in this study.

The spatial distributions of preferred speed illustrated in Figs. 6, 7A, and 8 were based on a nonquantal measurement of speed preference. The advantage of this approach was that it incorporated responses to all three FM speed conditions. However, it is possible that this measurement may have artificially conferred the status of a continuum on the distribution of FM speed-selective responses. On examination, the formula used for this analysis has an inherent bias that assigns a preference for medium FM sweeps to some responses that may not be speed selective at all. Yet, it is unlikely that this bias affected the essential aspects of the spatial distribution of preferred speed in any significant way for the following reasons. First, this bias only affected 10% of all the corti-

cal responses tested (i.e., 10% of all units in this study had no preferred speed as determined by the discrete measurement). Second, based on the nonquantal measurement, when the spatial distribution of preferred speed was compared with that based on the discrete speed criterion, a close agreement was observed between the distributions obtained with both estimates (see Fig. 7).

In general, the largest group of multiunit responses displaying a unimodal rate tuning profile, responded best to fast FM sweeps, the second largest group preferred medium FM sweeps, and the smallest group preferred slow FM sweeps. This finding is remarkable in light of the frequency transitions that have been measured in cat vocalizations (Pick 1979). It has been shown that approximately 90% of the transition rates of frequency glides in cat vocalizations are below 10 octaves per second. This rate corresponds to the "slow" FM speed used in the present study. Although Pick (1979) may not have chosen an exhaustive set of cat vocalizations for this analysis, the available data suggest that FM processing at the level of area AI of the cat cortex is not specialized for species-specific vocalizations. The range covered by the preferred FM speeds reported here and in the single-unit studies of Mendelson and colleagues (Mendelson and Cynader 1985; Mendelson and Grasse 1992) indicate that the cortex may be used for the detection of frequency transitions in the vocalizations of its prey (e.g., mice) or for the detection of spectral changes useful for sound localization. The fact that many cortical responses were sensitive to the direction of FM sweeps suggests that neurons were indeed responding to the rate of change of FM sweeps rather than to the integrated spectrum of frequency components covered by the FM sweep.

Direction selectivity

In general, all of the cases examined in the present study showed evidence of spatial segregation in the isofrequency domain for responses to either upward- or downward-directed FM sweeps. The dorsal region of the mapped areas contained units that, in general, responded best to upward-directed FM sweeps. In the central area of the maps, multiunits typically responded best to downward-directed FM sweeps. In most cases the preference in the ventral region was for upward-directed FM sweeps, though in some cases this region contained responses that were non-direction-selective or, at slow FM speeds, even showed some preference for downward-directed sweeps. For the most part, direction-selective responses appeared to be independent of CF. Variability between maps in the spatial distribution of preferred direction was most often observed in the ventral region of the cortex. Units located in the dorsal and middle regions of AI were, in general, more strongly direction-selective than units located in the ventral region. Thus, the degree of direction selectivity may be related to the variability observed. In addition, direction selectivity may also be related to the fact that significant correlation with integrated excitatory bandwidth was observed in the dorsal but not in the ventral region. Finally, as mentioned above for

preferred speed, it is possible that some of the variability observed for locations sampled in the ventral region are due to the inclusion of units belonging to AII rather than AI.

Several possible mechanisms underlying direction selectivity have been proposed by other investigators (e.g., Erulkar et al. 1968; Suga 1965a,b, 1968, 1969; Phillips et al. 1985; Shore and Nuttall 1985; Shore et al. 1987; Mendelson and Grasse 1992). With the exception of Mendelson and Grasse (1992), the suggested mechanisms are primarily monaural in nature. One prevalent type of monaural mechanism is characterized by inhibitory sidebands flanking one or both sides of the excitatory response area. Suga (1965a,b, 1968, 1969) and Watanabe (1972) have accounted for FM sweep direction selectivity in terms of the temporal sequence of the component frequencies in the sweep that successively activate inhibitory and excitatory response areas of neuronal response profiles. It has been suggested that asymmetric inhibitory sidebands can generate direction-selective responses to FM sweeps (Suga 1965a,b, 1968, 1969). Recently, Sutter and Schreiner (1991) have presented evidence that the characteristics of the upper and lower inhibitory sidebands of cortical response profiles change systematically across the dorsoventral extent of cat AI.

The spatial distribution of direction-selective responses observed in the present study may reflect underlying inhibitory organization similar to that observed by Sutter and Schreiner (1991). However, inhibitory sidebands may not completely account for direction-selective responses of cortical neurons. Vartanian (1974) has pointed out that cells with the same preference for direction of FM sweep can exhibit a wide variety of different tuning curves and different inhibitory areas. In addition, Mendelson and Grasse (1992) have recently provided evidence demonstrating that cortical direction selectivity need not arise exclusively from a monaural input but can also be elicited using *binaural* stimulation. This was most apparent in neurons that appeared non-direction-selective under monaural FM sweep conditions, yet were direction-selective when stimulated with binaural FM sweeps. For many of these binaural direction-selective responses, the basic underlying mechanism seemed to operate in a predominantly inhibitory fashion, while for other direction-selective neurons the binaural mechanism was excitatory.

Preferred speed versus preferred direction of FM sweep

In addition to calculating direction selectivity for the strongest response (independent of FM speed), direction selectivity was also determined independently for each of the three FM sweep speeds used in the present study and compared with that derived from the strongest response. In most cases, direction selectivity determined for fast and medium FM sweeps closely resembled direction selectivity determined for the strongest response, especially in the dorsal and middle regions. The greatest amount of variability of direction selectivity occurred with slow FM sweeps and was most apparent in the central and ventral

regions of the cortex. It is possible that much of this variability can be attributed to responses that were weaker than those obtained with medium and fast FM sweeps. The weaker responsivity to slow FM sweeps may be more susceptible than medium or fast FM sweep responses to random fluctuations. A reduced responsiveness for broadband stimuli has been observed in the central region of the dorsoventral extent of AI, coinciding with the areas of slow preferred speeds and downward-directed sweeps (Schreiner and Mendelson 1990). However, this coincidence alone cannot explain the relative loss of direction selectivity for slower speeds at the center of the maps, because those locations showed good direction selectivity for the faster FM speeds.

The smallest amount of variability in direction selectivity for the different speed conditions occurred in the dorsal region of AI. This suggests that direction selectivity of cells in dorsal AI is largely independent of FM speed. In contrast, for the central and particularly for the ventral aspect of AI, direction selectivity may be more dependent upon the speed of the FM sweep. However, the correlations of FM speed and direction for cells located in the dorsal region and for those located in the ventral region, were not statistically significant. Thus, while this may seem to contradict some of the results described (i.e., units in the mid-region of isofrequency contours appearing to prefer slow FM speeds and downward-directed sweeps while units located more dorsally prefer fast/ medium FM sweeps and upward-directed FM sweeps), the differences observed between dorsal and ventral regions may provide a clue as to the function of the multiple representations of upward-directed FM sweep responses and fast- and medium-selective speed responses distributed within AI. Further, while these relationships were not consistent across all subregions of AI, nor within nor between cases, the implication of these results is that FM speed may be processed independently of FM sweep direction in some but not all cells in AI.

Comparison with other FM studies

At the level of the auditory nerve, most neurons are not sensitive to FM sweep direction except at very fast speeds (Sinex and Geisler 1981). In the cochlear nucleus some researchers have found that cells prefer upward-directed FM sweeps at high rates of modulation (Britt and Starr 1976; Shore et al. 1987), while others have observed a preference for downward-directed FM sweeps, particularly when stimulus intensity is increased (Moller 1974). It has been suggested that in the inferior colliculus and medial geniculate body some neurons are selective for the direction and/or speed of FM sweeps (Watanabe 1972; Whitfield and Purser 1972; Clopton and Winfield 1974; Vartanian 1974; Poon et al. 1991). At the level of the cortex, where even greater integration of information from the periphery occurs, the majority of neurons have been shown to respond to the direction and speed of the FM sweeps as well as to the extent and form of the FM sweeps (Suga 1965a,b; Whitfield and Evans 1965; Nomoto 1980; Scheich and Bonke 1981; Mendelson and

Cynader 1985; Phillips et al. 1985; Mendelson and Grasse 1992).

Suga and his colleagues (Suga 1965a 1986; O'Neill and Suga 1982; Tsuzuki and Suga 1988; Edamatsu et al. 1989) were the first to demonstrate that responses to FM sweep stimuli are systematically distributed in the cortex of the bat. Responses to FM sweep parameters, such as types of FM sweeps (eg., FM_1-FM_2, FM_1-FM_3) and delay between the FM sweeps in a paired presentation are spatially distributed along two dimensions in a region of cortex devoted exclusively to processing FM stimuli. This is in contrast to the present study in which FM speed- and direction-selective responses appear to be spatially distributed along one cortical dimension. However, this apparent discrepancy may be readily explained by the fact that bats have a highly specialized auditory system equipped to handle their biosonar communication signals.

The present results are most directly comparable with the single-unit work of Mendelson and Cynader (1985) and Mendelson and Grasse (1992), as these studies employed similar FM sweeps to those used in the present study. In agreement with the present results, Mendelson and colleagues have found that most cortical neurons respond best to fast or medium FM sweeps. These authors have also reported that 59% of their cells (tested monaurally) in AI were direction selective. Of these direction-selective cells, the majority (74%) preferred downward-directed FM sweeps. In the present study, only 45% of the cortical locations sampled were direction selective. However, of these, almost two thirds of the direction-selective responses in the present study preferred downward-directed FM sweeps.

Several possible explanations may account for the difference in the percentage of direction-selective cells in the two types of studies. The first is that different types of FM sweeps were used: Mendelson and colleagues used linear FM sweeps (0.0–50.0 kHz), while the present mapping study used exponential FM sweeps (0.25–64.0 kHz). In pilot experiments leading up to the present mapping study, a comparison was made between responses to linear and exponential FM sweeps presented at two different FM speeds. It was found that some cells responded in a similar way to both linear and exponential FM sweeps, while other unit responses changed as a function of tl type of FM sweep presented. Thus, the type of FM sweep used may at least partially account for the difference between the single-unit and mapping studies.

A second explanation for the difference in the incidence of direction-selective responses may be related to the fact that the single-unit studies presented FM sweeps at approximately 20 dB above threshold, while the present mapping study presented FM sweeps at 40 dB above threshold. As mentioned above, one type of neural mechanism thought to account for direction selectivity is based on the presence of asymmetric inhibitory sidebands (Suga 1965a,b, 1968). In many cases these inhibitory bands only become apparent at higher intensity levels. Thus, it is possible that, in the present mapping study, more cells were encountered that possessed both lower and upper inhibitory sidebands, thereby rendering

the overall response of the unit non-direction-selective. Further support for the role of stimulus intensity can be derived from the work of Schreiner and Mendelson (1990). They found that while the cortical spatial distribution of Q_{10dB} and Q_{40dB} appeared similar, these two response properties were only weakly correlated, suggesting that at higher stimulus levels additional mechanisms may be involved which do not operate at lower levels. Thus, the direction selectivity observed at lower stimulus levels may not be manifest at higher stimulus levels.

Finally, the difference in the percentage of direction-selective cells between the two types of studies may also be due to single- versus multiunit recording. As mentioned earlier, a multiunit recording samples several neurons within a small area. It is possible that within a multiunit cluster there are some direction-selective neurons and some non-direction-selective neurons. Depending upon the number of direction and non-direction-selective neurons within a given cluster, it is possible that the non-direction-selective cells may dominate and diminish the effects of the direction-selective neurons, and vice versa. It is also conceivable that some units within a cluster prefer upward-directed FM sweeps while others prefer downward-directed FM sweeps, thus causing the net response to be non-direction selective. In this sense, the direction-selective responses at a given cortical location reflect the dominant neural activity characteristic of the cortical processing for that region.

Comparison with other auditory mapping studies

Distributions of other stimulus parameters across the isofrequency domain of AI have been discussed in detail previously (Schreiner and Mendelson 1990; Sutter and Schreiner 1991; Schreiner et al. 1992) and thus will only be treated summarily here. It is well known that the cortex is tonotopically organized, with cells that possess high CFs encountered in the anterior region and lower CFs in the posterior region of AI. Dorsoventrally, neurons with similar CFs are encountered along isofrequency contours (Rose and Woolsey 1948; Merzenich et al. 1975). One response feature that appears to be distributed orthogonally to CF is the type of binaural response. Imig and Adrian (1977) and, later, Middlebrooks et al. (1980) found alternating bands or elongated patches of EE and EI cells oriented approximately orthogonal to isofrequency contours. These results have recently been extended by Reale and Kettner (1986) who compared interaural intensity differences with equal-intensity stimuli. They found that the distribution of EE and EI cells in the dorsal and middle regions of AI was similar to that observed in other studies where equal-intensity stimuli had been used. However, in ventral AI, Reale and Kettner (1986) described a disproportionate number of cells with both facilitatory and inhibitory response properties. This segregation in binaural responses between the ventral and more dorsal regions of AI provides additional support for the present findings that ventral AI, excluding those sites that may belong to AII, may be functionally distinct from the more dorsal regions of AI.

Schreiner and Mendelson (1990) examined the spatial distribution of response properties to pure tone and broadband transient stimuli across the entire dorsoventral extent of AI. The most significant finding was the presence of a systematic distribution of the integrated excitatory bandwidth along the isofrequency domain. For Q_{10dB} and Q_{40dB}, the highest values were found in the middle of AI, while locations with lower Q values were located in more dorsal and ventral subregions, supporting earlier observations by Schreiner and Cynader (1984). Responses to broadband transients were also systematically distributed across the cortex such that those units most responsive to clicks were located in the more dorsal and ventral subregions.

In the present study it was found that the distribution of preferred speed and direction responses for FM sweeps were correlated with integrated excitatory bandwidth and broadband transient responses. A number of functional relationships were revealed by this analysis. While some of these statistically significant correlations were relatively weak, they still suggest that functional interactions may arise between these response parameters. These relatively low correlation values also suggest the presence of other factors, such as inhibitory sidebands, that are most likely involved in the neural mechanisms underlying speed and direction selectivity to FM sweeps. As described above, in the dorsal and ventral aspects of AI, multiunits preferred fast or medium upward-directed FM sweeps. Coincident with this FM sweep selectivity is that these cortical responses also tended to have low Q_{10dB} and Q_{40dB} values and responded best to broadband stimuli (Schreiner and Mendelson 1990). This stands in marked contrast to responses recorded in the middle region of the isofrequency domain that preferred slower (i.e., medium and/or slow FM sweeps) and downward-directed FM sweeps characterized by narrow tuning curves and poor responses to transients (Schreiner and Mendelson 1990).

Excitatory bandwidth and speed selectivity of a cortical location can be linked by the concept of *integration time*. A strong response is only elicited if the stimulus remains present within the cell's excitatory response area for some minimal period of time. Thus, slow FM speeds are required to excite neurons with narrow tuning curves, whereas faster speeds are sufficient for more broadly tuned neurons. The apparent link between excitatory bandwidth and direction selectivity (e.g., sites with narrow tuning curves responding preferentially to downward-directed sweeps) is less straight forward. In addition, it may also involve inhibitory sidebands that are probably used to sharpen (i.e., tighten) the bandwidth of neural tuning curves.

The significant correlation between responses to broadband transient stimuli and preferred FM speed is not unexpected given that a fast FM sweep can, in essence, be considered analogous to a brief broadband signal not unlike a click. Thus, those units that responded best to medium or fast FM sweeps also responded well to transient signals. The fact that there was no significant correlation between a response to a broadband transient signal and a direction-selective response for FM sweeps

suggests that there are different mechanisms for direction selectivity and preferred speed of FM sweeps. Additional support for this suggestion is derived from the lack of significant correlations between preferred speed and direction of FM sweep across the cortex.

Psychophysical studies

A number of investigators have provided evidence suggesting that FM and AM stimuli are processed independently by parallel pathways or "channels" in the auditory system (Kay and Matthews 1972; Gardner and Wilson 1979; Regan and Tansley 1979; Tansley and Regan 1979). A recent neuromagnetic study by Makela et al. (1987) has confirmed these findings in the human auditory cortex. In addition to the independent processing of AM and FM stimuli, several of these researchers have shown that upward-directed linear FM sweeps appear to be processed independently from downward-directed FM sweeps. Further, it has been suggested that channels exist in the auditory pathway that are sensitive to different rates of frequency modulation (Kay 1982). The global cortical mechanism described here is composed of orderly arrangements of responses to the direction and speed of FM sweeps and may help in understanding some of these psychophysical findings by maintaining close physiological parallels with these psychophysical direction- and rate-independent FM channels.

Conclusions

The present study shows that multiunit responses to the direction and speed of FM sweeps are spatially distributed in a nonrandom fashion along the dorsoventral axis of AI of the cat. In general, the pattern of the spatial distribution of responses to the speed of FM sweeps can be summarized as: fast-medium-slow-medium-fast, proceeding in a dorsal to ventral direction across AI. Furthermore, multiunit responses preferentially selective for upward-directed FM sweeps were located in the dorsal and ventral aspects of AI while multiunit responses preferring downward-directed FM sweeps were located in the mid-region of the isofrequency domain. In addition to the already-established tonotopicity and binaural response properties, these results provide further evidence for a systematic distribution of FM responses which defines a dimension of functional organization oriented along the isofrequency domain of AI. Finally, these results may provide evidence for a physiological basis for independent psychophysical channels underlying the perception of FM sweeps.

References

Britt R, Starr A (1976) Synaptic events and discharge patterns of cochlear nucleus. II. Frequency modulated tones. J Neurophysiol 31:179–193

Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. J Neurophysiol 36:1138–1158

Clopton BM, Winfield JA (1974) Unit responses in the inferior colliculus of rat to temporal auditory patterns of tone sweeps and noise bursts. Exp Neurol 42:532–540

Edamatsu H, Kawasaki M, Suga N (1989) Distribution of combination-sensitive neurons in the ventral fringe area of the auditory cortex of the mustached bat. J Neurophysiol 61:202–207

Erulkar SD, Butler RA, Gerstein GL (1968) Excitation and inhibition in cochlear nucleus. II. Frequency-modulated tones. J Neurophysiol 31:637–648

Gardner RB, Wilson JP (1979) Evidence for direction-specific channels in the processing of frequency modulation. J Acoust Soc Am 66:704–709

Hubel DH, Wiesel TN (1965) Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. J Neurophysiol 28:229–289

Imig TJ, Adrian HO (1977) Binaural columns in the primary field (AI) of auditory cortex. Brain Res 138:241–257

Imig TJ, Irons WA, Samson FR (1990) Single-unit selectivity to azimuthal direction and sound pressure level of noise burst in cat high-frequency primary auditory cortex. J Neurophysiol. 63:1448–1466

Kay RH (1982) Hearing of modulation in sounds. Physiol Rev 62:894–975

Kay RH, Matthews DR (1972) On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones. J Physiol (Lond) 225:657–677

Makela JP, Hari R, Linnankivi A (1987) Different analysis of frequency and amplitude modulations of a continuous tone in the human auditory cortex: a neuromagnetic study. Hear Res 27:257–264

Mendelson JR, Cynader MS (1985) Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation. Brain Res 327:331–335

Mendelson JR, Grasse KL (1992) FM sweep selectivity in cat primary auditory cortex: a comparison of monaural and binaural responses in single units. Exp Brain Res 91:435–454

Merzenich MM, Knight PL, Roth GL (1975) Representation of cochlea within primary auditory cortex in the cat. J Neurophysiol 38:231–249

Middlebrooks JC, Dykes RW, Merzenich MM (1980) Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. Brain Res 181:31–48

Moller AR (1974) Coding of sounds with rapidly varying spectrum in the cochlear nucleus. J Acoust Soc Am 55:631–640

Mountcastle VB (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. J Neurophysiol 20:408–434

Mountcastle VB (1978) In: Edelmann GM, Mountcastle VB (eds) The mindful brain: cortical organization and the group selective theory of higher brain function. MIT, Cambridge

Nomoto M (1980) Discharge patterns of the primary auditory cortex in cat. Jpn J Physiol 30:427–442

O'Neill WE, Suga N (1982) Encoding of target range information and its representation in the auditory cortex of the mustached bat. J Neurosci 47:225–255

Phillips DP, Hall SE (1987) Responses of single neurons in cat auditory cortex to time-varying stimuli: Linear amplitude modulations. Exp Brain Res 67:479–492

Phillips DP, Mendelson J, Cynader MS, Douglas RM (1985) Responses of single neurons in cat auditory cortex to time-varying stimuli: frequency-modulated tones of narrow excursion. Exp Brain Res 58:443–454

Pick GF (1979) A study of frequency transition in cat vocalization. J Acoust Soc Am 65:594–597

Poon PWF, Chen X, Hwang JC (1991) Basic determinants for FM responses in the inferior colliculus of rats. Exp Brain Res 83:598–606

Rajan R, Aitkin LM, Irvine DRF (1990) Azimuthal sensitivity of neurons in primary auditory cortex of cats. II Organization along frequency-band strips. J Neurophysiol. 64:888–902

Reale RA, Imig TJ (1980) Tonotopic organization in cat auditory cortex. J Comp Neurol 192:265–292

Reale RA, Kettner RE (1986) Topography of binaural organization in primary auditory cortex of the cat: effects of changing interaural intensity. J Neurophysiol 56:663–682

Rees A, Moller AR (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. Hear Res 10:301–330

Regan D, Tansley BW (1979) Selective adaptation to frequency-modulated tones: Evidence for an information-processing channel selectively sensitive to frequency changes. J Acoust Soc Am 65:1249–1257

Rose JE, Woolsey CN (1948) Structure and relations of limbic cortex and anterior thalamic nuclei in rabbit and cat. J Comp Neurol 89:279–348

Scheich H, Bonke BA (1981) Tone-versus FM- induced patterns of excitation and suppression in the 14- C-2-deoxyglucose labelled auditory "cortex" of the Guinea Fowl. Exp Br Res 44:445–449

Schreiner CE, Cynader MS (1984) Basic functional organization of second auditory cortical field (AII) of the cat. J Neurophysiol 51:1284–1305

Schreiner CE, Langner G (1988) Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. J Neurophysiol 60:1823–1840

Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. J Neurophysiol 64:1442–1459

Schreiner CE, Urbas JV (1986) Representation of amplitude in the auditory cortex of the cat. I. The anterior auditory field (AAF). Hear Res 21:227–241

Schreiner CE, Urbas JV (1988) Representation of amplitude in the auditory cortex of the cat. II. Comparison between cortical fields. Hear Res 32:49–64

Schreiner CE, Mendelson JR, Sutter ML (1992) Functional topography of cat primary auditory cortex: representation of tone intensity. Exp Brain Res 92:105–122

Shore SE, Nuttall AL (1985) High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. J Acoust Soc Am 78:1286–1295

Shore SE, Clopton BM, Au YN (1987) Unit responses in ventral cochlear nucleus reflect cochlear coding of rapid frequency-sweeps. J Acoust Soc Am 82:471–478

Sinex DG, Geisler CD (1981) Auditory-nerve fiber responses to frequency-modulated tones. Hear Res 4:127–148

Suga N (1965a) Functional properties of auditory neurones in the cortex of echo-locating bats. J Physiol 181:671–700

Suga N (1965b) Analysis of frequency modulated sounds by auditory neurons of echo-locating bats. J Physiol (Lond) 179:26–53

Suga N (1968) Analysis of frequency-modulated and complex sounds by single auditory neurones of bats. J Physiol (Lond) 198:51–80

Suga N (1969) Classification of inferior collicular neurones of bats in terms of responses to pure tones, FM sounds, and noise bursts. J Physiol (Lond) 200:555–574

Sutter M, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. J Neurophysiol 65:1207–1226

Tansley BW, Regan D (1979) Separate auditory channels for unidirectional frequency modulation and unidirectional amplitude modulation. Sensory Proc 3:132–140

Tsuzuki K, Suga N (1988) Combination-sensitive neurons in the ventroanterior area of the auditory cortex of the mustached bat. J Neurophysiol 60:1908–1923

Vartanian IA (1974) On mechanisms of specialized reactions of central auditory neurons to frequency-modulated sounds. Acustica 31:305–310

Watanabe T (1972) Fundamental study of the neural mechanism in cats subserving the feature extraction process of complex sounds. Jpn J Physiol 22:569–583

Whitfield IC, Evans EF (1965) Responses of auditory cortical neurons to stimuli of changing frequency. J Neurophysiol 28:655–672

Whitfield IC, Purser D (1972) Microelectrode study of the medial geniculate body in unanaesthetized free-moving cats. Brain Behav Evol 6:311–322

Woolsey CN, Walzl EM (1942) Topical projection of nerve fibers from local regions of the cochlea in the cerebral cortex of the cat. Bull Johns Hopkins Hosp 71:315–344

Functional Topography of Cat Primary Auditory Cortex:
Response Latencies and Binaural Classes

Julie R. Mendelson, Christoph E. Schreiner. Mitchell L. Sutter

Address to which correspondence should be sent:

C.E. Schreiner

Coleman Laboratory, Keck Center for Integrative Neuroscience, Box 0732, University of California at San Francisco, San Francisco, CA 94143-0732, USA

J.R. Mendelson

Division of Life Sciences, University of Toronto, Scarborough College, Scarborough Ontario, M1C 7S7, Canada

M.L. Sutter

Center of Neuroscience, University of California at Davis, 1544 Newton Crt., Davis, CA 95616, USA

Abstract

The neuronal responses to tones, broadband transient stimuli, and frequency-modulated (FM) sweeps were mapped in the primary auditory cortex (AI) of barbiturate-anaesthetized cats. The spatial distribution of the final two response parameters in this series of experiments, onset latency and binaural response, were determined in four cases. The functional relationship between these two parameters and the previously reported spectral (Q_{10dB}, Q_{40dB}, broadband transient stimulus), intensity (threshold, strongest response level (SRL), dynamic range and monotonicity) and temporal (FM sweep direction and speed) parameters was subjected to both a global and regional analysis. Onset latency responses were systematically distributed along the dorsoventral/isofrequency axis of AI such that units with shorter latencies were located in the central region while units with longer latencies were more often found in the dorsal and ventral portions of the cortex. As shown by other investigators, alternating bands or patches of EE and EI units were also distributed across the cortex. The results of a point-by-point analysis revealed a lack of dependency of onset latency on binaural response type. However, when the relationship between onset latency response and the remaining 10 response parameters was examined, the correlations with the following parameters were found to be statistically significant: Q_{10dB}, Q_{40dB}, broadband transient response, strongest response level (SRL), monotonicity, and preferred FM sweep direction. These correlations suggest that units that respond with shorter onset latencies are more sharply tuned, do not respond well to broadband stimuli, have a higher degree of nonmonotonicity, and prefer FM sweeps that change from a high to a low frequency. Binaural response type was found to be significantly correlated with Q_{40dB}, broadband transient response, threshold, SRL, dynamic range, and preferred FM sweep speed. These correlations imply that EI neurons tend to have sharper tuning, weaker responses to broadband stimuli, lower thresholds, lower SRLs, a narrower dynamic range, higher degree of nonmonotonicity, and prefer slower FM sweeps than EE units. The systematic distribution and the functional relationship between these response parameters may provide the representational basis for the detection and identification of specific features in the animal's natural environment.

Keywords: Primary auditory cortex - Onset latency - Binaural response type - Topographical functional organization - Cat

2

Introduction

Previous investigations of the functional organization of the primary auditory cortex (AI) of the cat have revealed that a number of response properties appear to be systematically distributed within AI. One of the most prominent and best documented of these response properties is the representation of the characteristic frequency (CF) of neurons (e.g., Woolsey and Walzl, 1942; Merzenich et al., 1975; Reale and Imig, 1980). This property is revealed as an increase in the CF of neurons along the rostrocaudal extent of AI. A second prominent organizational feature of AI is the spatial segregation of binaural facilitatory (EE) and binaural inhibitory (EI) cells along the dorsoventral extent of AI (Imig and Adrian, 1977; Middlebrooks et al., 1980). Reale and Kettner (1986) have extended these findings by showing that cells located in the ventral region of cat AI contain both binaural facilitatory and inhibitory response characteristics, depending on the interaural intensity differences of the stimuli presented. Recently, a number of investigators have observed that other response properties, such as spectral characteristics, inhibitory sidebands, frequency modulated (FM) sweep selectivity, as well as a subset of sound localization cues, appear to be distributed in a non-random fashion within AI (Middlebrooks and Pettigrew, 1981; Jenkins and Merzenich, 1984; Reale and Kettner, 1986; Imig et al, 1990; Rajan et al, 1990; Schreiner and Mendelson, 1990; Sutter and Schreiner, 1991; Heil et al., 1992; Mendelson et al., 1993; Shamma et al., 1993).

In a recent series of studies, we examined the spatial distribution of twelve properties of single- and multiple-unit response profiles in AI. One of the goals of these experiments was to examine a relatively large number of response properties at individual cortical locations in order to compare different response parameter distributions directly and to ascertain their potential interrelationships. In the first report in this series of experiments (Schreiner and Mendelson, 1990), we demonstrated that the "integrated excitatory bandwidth" (i.e., the locally effective frequency range that produces excitation), as well as responses to broadband transients, were topographically arranged within AI. The principal gradient of the integrated excitatory bandwidth was found to be oriented orthogonally to the cochleotopic representation such that narrowly tuned cortical neurons were typically located in the derso-ventral center of AI while more broadly tuned units were found in more dorsal and ventral subregions of the cortex. In addition, we

3

found that cortical locations in which cells displayed strong responses to broadband transient stimuli generally favoured the extreme dorsal and ventral subregions of AI and not the sharply tuned central subregion. Additional response parameters such as multi-peaked tuning curves, inhibitory sidebands (Sutter and Schreiner, 1991), stimulus intensity responses (eg., threshold, strongest response level, dynamic range) and monotonicity (Schreiner et al. 1992) have also been found in these studies to be non-uniformly distributed along the dorsoventral extent of AI. Finally, cortical responses to the direction (changing from a low to a high frequency or vice versa) and speed (rate of change of frequency over time) of FM sweeps have been shown to exhibit an organized spatial distribution along this isofrequency axis of AI in cat (Mendelson et al , 1993) and ferret (Shamma et al., 1993).

The present report describes the spatial distribution of two additional response properties in this series of experiments, onset latency and binaural interaction responses, and compares their characteristics to those previously described. Onset latency is a fundamental descriptor of neuronal responses and has been studied in single units in the cortex (Phillips and Irvine, 1981; Phillips et al., 1985). However, the transcortical spatial distribution of these responses and the relationship of onset timing behavior to other response parameters has not been thoroughly investigated. While the spatial distribution of binaural responses has been examined extensively, its inclusion in the present paper provides a more complete and detailed investigation of the functional topography of twelve different response parameters in AI studied under identical conditions. In this way, an in-depth analysis of the covariance of all twelve response parameters can be carried out in parallel. A subset of these correlations will be presented in the present paper while a global comparison between all parameters investigated thus far will be dealt with in a subsequent paper in this series.

Materials and Methods

Preparation

A detailed description of the preparation and recording methods has been provided in a previous paper in this series (see Schreiner and Mendelson, 1990) and thus, will only be summarized here. Extensive spatial mapping of multiple unit responses was conducted on four adult cats. In addition, the

latency distribution of single unit responses along the isofrequency domain was obtained (see Schreiner and Sutter, 1992; Sutter and Schreiner 1994). Initially, animals were tranquilized with an intramuscular injection of ketamine hydrochloride (10 mg/kg) and acetylpromazine maleate (0.1 mg/kg) to allow for venous cannulation. Animals were then given sodium pentobarbital (30 mg/kg i.v.) followed by intramuscular injections of atropine (1 mg im) to reduce salivation and dexamethasone sodium phosphate (0.14 mg/kg im) to prevent brain edema. A surgical level of anaesthesia was maintained throughout the experiment with a constant i.v. infusion of sodium pentobarbital (2 mg/kg/hr) in lactated Ringer's solution. Body temperature was maintained at 37.5 C.

Both pinnae were surgically reflected and the external meatuses exposed to allow for insertion of speaker tubes within 3 mm of the tympanic membranes. The temporal muscle was reflected on the right side and a craniotomy was performed over the middle of the ectosylvian gyrus. The dura mater was removed and the brain was bathed in silicone oil. A photograph of the surface vasculature was used to record the electrode penetration sites.

Recording procedure

The animal was placed in a modified head-holder that allowed unobstructed access to the external meati. Recording took place in an electrically-shielded sound-attenuating chamber (IAC). Parylene-insulated tungsten electrodes were introduced into the cortex orthogonal to the surface (as viewed through a Zeiss operating microscope) by means of a remote-controlled stepping motor microdrive (Kopf). Single unit and unit cluster responses were examined at one location per electrode penetration at a depth between 600 and 1000 μm below the cortical surface, corresponding approximately to layers III and IV. The number of units per cell cluster was estimated to be between 2 and 6. Stimulus-evoked action potential event times were collected and stored on-line by a PDP 11/73 computer.

In all cases, a cortical surface area of approximately 3 x 5 mm^2 to 3 x 15 mm^2 was examined with 80-150 electrode penetrations in each map. The recording strategy was to map the isofrequency domain of AI from its most dorsal to its most ventral aspects, often extending into the second auditory field (AII). The range of isofrequency contours investigated across all animals ranged from 3.2 to 28.0 kHz.

5

Stimulus Generation

Tone burst stimuli, generated by a microprocessor, (TMS32010; 16 bit D/A converter at 120 kHz; low-pass filter at 35 kHz) were 50 ms in duration with a 3 ms rise/fall time. The interstimulus interval was 500-1000 ms. Tone bursts were presented in a pseudo-random sequence of different frequency/level combinations selected from 15 level and 45 frequency values (geometrically centred around the CF). Stimulus levels were in 5 dB increments, resulting in a dynamic range of 70 dB. Responses to a minimum of 675 different tone bursts were recorded at each electrode site. In addition to the pure tone burst stimuli, animals were also presented with broadband transient signals (see Schreiner and Mendelson, 1990) and exponential FM sweeps (see Mendelson et al. 1993). FM sweeps extended from 0.25 to 64.0 kHz (upward-directed) or from 64.0 to 0.25 kHz (downward-directed). The rates of frequency change (i.e., sweep speeds) used were 110 octaves/sec, 33 octaves/sec, and 10 octaves/sec with corresponding sweep durations of 72 ms, 240 ms, and 800 ms, respectively.

All stimuli were attenuated (Hewlett-Packard) and passed through transducers (STAX 54) that were attached to sound delivery tubes sealed into the auditory meati. The stimulus delivery system was calibrated for pure tones with a sound level meter (Bruel & Kjaer 2209). The transfer function of the sound delivery system showed a roll-off of 10 dB/oct above 14 kHz and additional deviations of up to 6 dB were possible (see Schreiner and Mendelson, 1990).

Data Analysis

Frequency response areas (FRAs) for monaural stimulation (contralateral ear) were objectively determined from the 675 frequency/level combinations recorded at each electrode site. From this the integrated excitatory bandwidth at 10 and 40 dB (Q_{10dB} and Q_{40dB}, respectively) above threshold were derived. Shortest (i.e., minimum) onset latency and latency at 30 to 40 dB above minimum threshold were ascertained from the latency/level function determined for each unit or unit cluster. The arrival time of the first spike of the response that was consistent with those obtained for neighboring stimulus intensities was taken as the onset latency measure. Binaural interaction responses were assessed by presenting CF to both

6

ears at equal intensity, usually 5 to 15 dB above minimum threshold. Multiple-unit or single unit responses were then categorized as binaural facilitatory (EE), binaural inhibitory (EI), or binaural occluder (EO) responses (whereby the ipsilateral ear had no significant effect on the binaural response of the unit). The strength of the unit's response to broadband transient stimuli was also determined at each cortical location (expressed in % of the maximum response to a pure tone stimulus). In addition, the following response properties were assessed: contralateral threshold, monotonicity, dynamic range, and strongest response level (SRL) preferred speed and preferred direction of FM sweep (see Schreiner et al., 1992; Mendelson et al., 1993).

Pseudo three-dimensional projections and contour plots were utilized to represent the spatial distribution of response parameters across the cortical surface (Golden Software, Inc.). The cortical locations of the recording sites were used to generate a two-dimensional grid of the represented area by projecting the actual sites to the nearest grid point. Elevation of the gridded surface corresponds to the local spatially averaged magnitude of a functional parameter (e.g., onset latency) at a given site. The elevated surfaces are represented either as the orthographic projection of the surface grid or of the elevated contours of equal parameter values (e.g., isolatency or isofrequency contours). In some cases, the actual onset latencies are indicated in a box below the pseudo three-dimensional projection for direct comparison of the value and location of each measurement. The interpolation algorithm used (inverse distance squared) inherently applied some spatial smoothing to the data, resulting mainly in a slight compression of the encountered functional range and a local averaging of closely spaced adjacent points (see Schreiner and Mendelson, 1990). This method emphasizes, although only slightly, topographical organizations with low spatial frequency and de-emphasizes steep local gradients or variabilities. It should be kept in mind that not every grid point is directly supported by a datum point (equivalently, not every section in a linear or polynomial fit of a two-dimensional data distribution is equally supported by data points). The 'smoothing' of the surface plots represents only a small distortion of the data that is not unlike a two-point or three-point averaging in a two-dimensional data representation. In this case, the interpolation can be considered to occur along both sets of gridlines.

Results

A total of 463 single and multiple unit responses for which a CF could be determined were recorded from four extensively mapped animals. In addition, 104 single units and 111 multiple units encountered along empirically determined isofrequency axes were investigated in four additional animals. For multiple-unit mapping, the average distance between each recording site was approximately 150-350 μm. Within each map, a large portion (>75%) of the tonotopic map indicative of AI was observed (Merzenich et al., 1975; Reale and Imig, 1980). For three of the four cases, the spatial distributions of characteristic frequency (CF), integrated excitatory bandwidth (expressed as Q_{10dB} or Q_{40dB}), broadband transient stimulus (Schreiner and Mendelson, 1990), FM sweep responses (Mendelson et al., 1993), and tone intensity responses (Schreiner et al., 1992) are described in the other papers in this series and thus will only be briefly summarized here. We begin with mapping data (N= 186) for one case which has not been previously described and for which the spatial distribution of the frequency tuning curve (or CF), Q_{10dB}, and minimum threshold (case #86-173) have been determined.

The CF of units for case #86-173 increased monotonically from approximately 4 to 13 kHz (Fig. 1A) as the electrode was moved in a caudal to rostral direction for most parts of the cortical area mapped. In the ventral portion of the map, a sudden increase in CFs above 16 kHz indicates the transition to the second auditory field (AII; Schreiner and Cynader, 1984). The integrated excitatory bandwidth of the multi-unit responses for this case changed systematically along the dorsoventral axis of area AI and across its ventral border with area AII (Fig. 1B). The ventral portion of the mapped area contained units that possessed the lowest Q_{10dB} values corresponding to broad integrated excitatory bandwidths. Many of these ventral multiple-unit responses were most likely located in AII (Schreiner and Cynader, 1984). Q_{10dB} values increased continuously toward the center of AI, with the narrowest bandwidths recorded approximately 2 mm dorsal to the AI/AII border. This central region with narrow integrated excitatory bandwidths and a strict cochleotopic gradient is indicative of the classically defined primary auditory area (Merzenich et al., 1975; Reale and Imig, 1980). Moving from central AI toward the suprasylvian sulcus, bandwidth sharpness gradually decreased. In the most dorsal portion of the mapped region, the principal cochleotopic gradient appeared to be maintained despite a slight change in the orientation of the

8

isofrequency contours (see Fig. 1A) that is consistent with previous reports (Middlebrooks and Zook, 1983; Sutter and Schreiner, 1992).

In addition to the two-fold representation of the integrated excitatory bandwidth falling along the dorso-ventral axis of AI, a second, less well-defined gradient of tuning bandwidth was evident in the central portion of AI running along the postero-anterior axis. This second gradient was in register with the frequency gradient typically observed in single unit data (Phillips and Irvine, 1980). The basic features of the distribution of integrated excitatory bandwidth across the middle ectosylvian gyrus revealed by case #86-173 was, in general, consistent with those found in all other cases studied thus far (see Schreiner and Mendelson, 1990).

The spatial distribution of minimum response thresholds for case #86-173 is shown in Figure 1C. A low threshold region was evident in the central and dorsal subregions with higher thresholds in the ventral and most dorsal portions of the map. An increase of threshold along the cochleotopic gradient toward higher CFs was also observed. A broad region of lowest thresholds was found to be in close alignment with the region of overall sharpest tuning (see Schreiner et al., 1992; Heil et al., 1992).

Representation of Onset Latency

On average, onset latency responses to CF tones presented 30 -40 dB above threshold varied between 8 and 18 ms for all cases studied. Fig. 2 illustrates representative latency/intensity level profiles for four multiple-units from case #87-518. Latency was measured in 5 dB steps over a range of 70 dB. As can be seen in all four panels, onset latencies were longer at lower intensity levels. For the response profile illustrated in Fig. 2a, onset latency decreased non-linearly as stimulus intensity increased. In Fig. 2b, there was a precipitous decrease in the onset latency between 10 - 25 dB at which point latency reached asymptote. The onset latency profile shown in Fig. 2c gradually decreased as stimulus intensity increased until approximately 60 dB above threshold where the onset latency began to increase. Fig. 2d shows a unit for which the onset latency decreased gradually and did not asymptote until 40 dB above threshold. All other recorded latency/level profiles resembled one of these examples.

9

Fig. 3 illustrates the relationship between onset latency and CF for each of the four intensively studied cases. As can be seen, the large variance in the onset latency data obscures any potential dependency upon CF. In other words, a wide range of onset latencies were encountered for each narrow range of CFs. The overall increase of latency for CFs below 15kHz in case #87-001 (Fig. 2B) is not necessarily a reflection of the latency/CF relationship suggested from studies of the peripheral auditory system (eg. Møller, 1977). If one postulates a non-uniform spatial distribution of latency values along the dorsal-ventral extent of AI, this increase in latency might just be a reflection of an incomplete spatial sampling of the isofrequency domain. Interestingly, the orientation of the mapped cortical area of cat #87-001 was turned 24° with regard to the orientation of the isofrequency domain (see Fig. 10; see also Fig. 4 in Schreiner and Mendelson, 1990). Therefore it is conceivable that a short-latency region of the lowest CFs for this case was missed during the mapping.

Overall, the global spatial distribution of onset latency values obtained at 30- 40 dB above threshold exhibited a non-random organization. There was a tendency for multiple-units located in the more dorsal and ventral subregions of AI, as well as those located in dorsal AII, to respond with longer onset latencies than those encountered in central AI. In some cases, variability in onset latency was observed in the dorsal portion of AI. Fig. 4 illustrates an example (#87-001) of the spatially reconstructed distribution of onset latencies observed over a high-frequency zone of AI. As mentioned above for this case, units with CF's between 10.3 and 26.3 kHz were examined. Below the pseudo-three-dimensional plot is an insert showing the latency values (in msec) to CF observed at each cortical location. In the pseudo-three-dimensional projections, the arrowheads along the X-axis indicate the approximate orientation of the isofrequency contours. For this case the orientation of the mapped area was tilted 24° relative to the isofrequency orientation, resulting in an approximate alignment of the 18.5 kHz contour with the ventro-posterior to dorso-anterior axis of the map. The Y-axis represents the dorsoventral extent over 5 mm. The elevation of the surface of the plot (i.e., along the Z-axis) indicates the onset latencies at the recording locations. Low elevations indicate shorter onset latencies, while higher elevations indicate longer onset latencies In general, longer onset latencies were observed in the dorsal subregion of AI and to a lesser extent, in the ventral subregion. This is apparent when considering the isofrequency contours

10

between 18.5 and 22.0 kHz. At lower frequencies, longer onset latencies were observed, especially in the dorsal and central aspects of AI.

Fig. 5 shows pseudo-three-dimensional plots of the onset latency distribution for three other cases. Fig. 5a shows the spatial distribution for #86-173. As can be seen, longer onset response latencies were typically encountered in the more dorsal and ventral aspects of AI (including the dorsal aspect of AII, see Figure 1A). Shorter onset latencies were observed in the more central subregion of AI. The case illustrated in Fig 5b differed from the others in that three subregions were disclosed containing relatively longer onset latencies and two subregions containing shorter onset latencies. The spatial frequency of longer and shorter onset latencies for this case (i.e., spatial distance between maxima and minima) was higher than that observed in other cases. For the example illustrated in Fig. 5c the short onset latency zone occupied a relatively smaller area in the central portion of cortex than in other cases, while longer latencies were found over a larger and more ventral zone

A comparison of the spatial distributions of onset latencies for single and multi-units is shown in Figs. 6-8. These spatial distributions are depicted in relation to the map derived for integrated excitatory bandwidth calculated at 40 dB above threshold (or *BW40*, see Schreiner and Sutter, 1992; Sutter and Schreiner, 1994) In order to compare the spatial distributions of onset latencies across cases, the minimum in the BW40 distribution (i.e., the region where cells were most sharply tuned at 40 dB above threshold) was assigned the reference value of 0 mm The positions of dorsal (negative values) and ventral (positive values) recording points are shown relative to this 0 mm position A scatter plot for single unit (Fig. 6a) and multi-unit (Fig. 6b) latencies calculated at 30 dB above threshold is shown as a function of their dorsoventral location in the cortex relative to BW40 In general, onset latencies tended to be shorter for multi-unit responses than for single unit responses (MU 16.24±5 16, SU 19 31±7.03) In addition, particularly for units located in the central subregion of AI, there was a greater scatter of onset latencies for single units (Fig. 6a) as compared to multiple-units (Fig 6b) However, the overall pattern of the spatial distribution for single- and multiple-unit responses was similar with a higher proportion of shorter onset latencies located in the central subregion of AI This is illustrated in the bar histogram in Fig 7 where variance (indicated by the error bars) is greater for single (Fig 7a) than for multi-units (Fig 7b).

98

particulary in the central subregion of the cortex. Finally, there appeared to be smaller variance of onset latency responses in the central as opposed to the ventral and especially the dorsal portions of the cortex.

Fig. 8 shows the spatial distribution of the minimum latency for single and multiple-unit responses. Minimum latency was determined by noting the stimulus condition that elicited the shortest latency response independent of stimulus intensity. This point typically corresponded to the asymptote of the latency/intensity level curve. As with the latencies obtained at 30dB above threshold, minimum onset latencies for multi-units were shorter than for single units (MU: 13.09±3.88ms, SU: 15.04±4.28ms). This spatial distribution resembles the one derived for onset latency at 30 dB above threshold with the exception that latencies were shorter and there was greater variance in the spatial distribution calculated for minimum onset response

Representation of Binaural Response

Fig. 9 illustrates the spatial distributions of binaural responses for the four cases described in this paper. Fig. 9a shows an enlarged schematic drawing of the cortical area (shown in the insert on the right) investigated in case #86-173. The filled circles represent those locations where multi-unit responses were classified as EE, the open circles represent the locations of EI cells, and the filled triangles indicate EO cells. In general, EE cells were typically found in the dorsal subregion of AI. In the rest of the cortical area examined, alternating bands or patches of EE and EI responses were observed. Fig. 9b is derived from the data illustrated in Fig. 9a and is redrawn in the form of a contour plot. The dashed line represents the border between AI and AII as revealed by the integrated excitatory bandwidth measurements (Schreiner and Cynader, 1984; Schreiner and Mendelson, 1990). The stippled areas represent locations where EE or EO cells were encountered and the white or blank areas represent the locations of EI cells. For this case, the spatial representation of binaural responses was comprised largely of clusters or patches as opposed to alternating elongated bands of EE and EI units   contour plots of binaural responses for the other three cases are illustrated in Fig. 9c (#87-001), 9d (#87/-518), and 9e (#87-706), respectively. Although not all bands span the entire mapped regions, in these latter three cases, EE and EI responses showed a much

Onset Latency versus Binaural Response

An attempt was made to correlate the spatial distribution of binaural response type with onset latency. Table I shows the average onset latency response of the binaural classes for each case. Only case #86-173 contained a sufficient number of EO responses to be included in the analysis. Fig. 10 shows the relationship between these two response parameters for case #87-518 (Fig. 10a) and case #87-001 (Fig. 10b). Contour plots have been generated to illustrate this point more clearly. The contour plots in the left panel of each side of Figure 10 show the spatial distribution of binaural response while the plots in the right panel show the spatial distribution of onset latency. The dashed line represents a prominent isofrequency contour for each case (viz., 10.0 and 18.0 kHz, respectively). The shaded areas in the binaural response contour plots demarcate those subregions where EE or EO unit clusters were encountered while the shaded areas in the onset latency plots illustrate subregions where onset response latencies were less than 10.6 ms. In general, these figures support the observation that there is little covariance between onset latency and binaural responses. For case #87-001 (Fig. 10b), there does appear to be some dependency of onset latency response on binaural response type. However, as revealed by the results of a point-by-point correlational analysis (Table I) this observation was not statistically significant. A significant covariance between these two response parameters was only evident for one of the four cases; #87-706 ($p < 0.03$).

Onset latency versus spectral, intensity, and temporal response properties obtained at the same cortical locations

In an attempt to further our understanding of the functional topography of AI, the relationship between onset latency and response parameters previoulsy described (Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993) was examined in the present study. To this end, statistical correlations between these response parameters were analyzed over the entire cortical region (AI) in the

13

three cases for which complete data sets were obtained: 87-001, 87-518, 87-706. In addition, statistical correlations were also determined separately for cells located in the dorsal (AId) and the ventral subregions (AIv) of the cortex. Figure 11 illustrates how the isofrequency domain was divided based on its Q10-dB distribution. The central portion of AI was defined as that region that contained the majority of the sharply tuned cortical sites. AId represents recording locations in that region of the cortex dorsal to the most sharply tuned responses while AIv represents those units located ventral to it.

Spectral Bandwidth Parameters (Q_{10dB}, Q_{40dB}*, broadband transient response).* In a series of experiments examining the functional topography of auditory cortex, Schreiner and Mendelson (1990) showed that overall there was a fairly systematic distribution of the integrated excitatory bandwidth along the isofrequency domain. For Q_{10dB} and Q_{40dB}, the highest values for both were found in the central region of AI while locations with lower Q values were located in more dorsal and ventral subregions. Responses to broadband transient stimuli were also systematically distributed across the cortex such that those recording sites where units were most responsive to clicks were located in the more dorsal and ventral subregions of the cortex, coinciding with more broadly tuned regions.

The correlation of integrated excitatory bandwidth and broadband transient responses with onset latencies is shown in Table II. Only the statistically significant correlations are included. As can be seen, for two of the three cases (#87-106 and #87-001) there was a strong negative correlation between onset latency and Q_{10dB} across the entire mapped region of cortex. When examined by cortical subregion, both of these cases exhibited a statistically significant correlation between these two response parameters in the dorsal but not the ventral portion of the cortex. These results suggest that sharply tuned cells respond with shorter onset latencies. For Q_{40dB} there was a strong negative correlation with onset latency for these same two cases. However, unlike Q_{10dB}, the significant correlation with Q_{40dB} was observed across the entire region of AI as well as in the dorsal subregion for #87-001 while it was only statistically significant in the ventral subregion of the cortex for #87-106.

For onset latency and broadband transient response a statistically significant negative correlation was observed in all three cases. This suggest that cells which respond vigorously to a broadband stimulus respond with longer onset latencies. For two of the cases, the correlation was only evident when the entire

extent of the cortical region studied was considered. For the third case, the correlation was only significant in the ventral subregion of the cortex.

Tone Intensity Parameters (monotonicity, SRL, threshold, dynamic range). Schreiner et al. (1992) reported that unit clusters with monotonic rate/level functions were spatially segregated from locations with nonmonotonic rate/level functions. Typically, two noncontiguous areas with a high degree of nonmonotonicity were observed: one was located at the dorso-ventral center of AI and a second in the dorsal third of AI. The stimulus levels that produced the highest SRL at any sampled location were also found to be systematically distributed across the cortex such that a region of low SRLs was always found in the central portion of AI. The spatial distribution of contralateral monaural threshold indicated several segregated areas containing clusters with either higher or lower response thresholds. Finally, the dynamic range, defined as the monotonic, fast-growing, low-level portion of the rate/level function did not reveal a clear systematic spatial distribution that appeared to be aligned with either the CF gradient or the isofrequency axis.

The analysis of onset latency and tone intensity responses yielded several significant correlations. For two of the three cases (#87-518 and 87-001), there was a significant negative correlation between monotonicity and onset latency for the entire cortex examined. This suggests that units with monotonic rate/level functions tend to respond with shorter onset latencies and that units with high degrees of nonmonotonicity respond with slightly longer latencies. SRL was also found to be significantly correlated with onset latency for two of the three cases. For one of the cases (#87-001) the significant correlation was only apparent in the ventral subregion of AI while for the second case (#87-705) it was significant for the entire extent of cortex examined. Finally, onset latency was significantly correlated with two other tone intensity parameters: contralateral threshold for case #87-706 and dynamic range for case #87-001. For both of these correlations, statistical significance was only observed in the ventral subregion of the cortex.

Temporal Parameters (FM Sweeps). Mendelson et al. (1993) showed that preferred speed of FM sweep appeared to be systematically distributed along the dorsoventral axis of area AI. In the dorsal subregion, cortical cells typically preferred fast and/or medium FM sweeps, followed more ventrally by cells that preferred medium, then slow, then medium speed FM sweeps. In the more ventral aspect of AI

15

neurons generally preferred fast FM sweeps. Direction selectivity also appeared to be nonrandomly distributed along the dorsoventral axis of AI. Units that preferred upward-directed FM sweeps were typically located in the more dorsal and ventral aspects of AI while units that preferred downward-directed FM sweeps were located in the central portion of AI.

Comparison of the spatial distribution of onset latency and FM sweep responses revealed that there was a statistically significant correlation between latency and preferred direction of FM sweep. This correlation was significant for the entire extent of AI as well as for the dorsal subregion. For two of the cases (#87-706 and 87-001), there was a positive correlation while for the third case (#87-518) there was a negative correlation between these response parameters. In other words, for two of the cases (#87-706 and 87-001), cells preferring upward-directed FM sweeps tended to respond with longer onset latencies while cells preferring downward-directed sweeps tended to respond with shorter onset latencies. This relationship was reversed for case #87-518.

Binaural response type versus spectral, intensity, and temporal response properties obtained at the same cortical locations

In order to assess potential differences between the three types of binaural responses, an analysis of the data was performed across all three of the more extensively mapped cases irrespective of spatial location. The fourth case (#86-173) is not considered for this analysis, since only three out of ten parameters were available for this case. The global means and standard deviation for these three cases are provided in Table III. For this analysis, significant differences ($p<0.05$) between EE and EI neurons were observed for 6 of the 10 tested response parameters: Q_{40dB}, broadband transient response, minimum threshold, SRL, dynamic range, and preferred FM speed. EI neurons tended to have sharper tuning, weaker broadband responses, lower thresholds and SRL smaller dynamic range, and prefereed slower FM speeds than EE neurons. Furthermore, some differences between EO responses and EI responses were also observed. However, the small number of EO responses for these three animals combined (N = 15; less than half of the omitted case) may be responsible for the lack of statistically significant differences between EO and EE/EI for most of the 10 parameters.

16

The data were also subjected to a subregional analysis in which the dorso-ventral axis of AI was divided into three subregions; the dorsal, central, and ventral portions, based on the sharpness of tuning distribution (see also Schreiner and Sutter, 1992). The analysis of differences between binaural response types was then performed for each subregion separately. Table IV illustrates the results of the binaural analysis for all three animals based on the three subregions.

Spectral Bandwidth Parameters (Q_{10dB} Q_{40dB}, *broadband transient response*). In general, there was a weak relationship between the spatial distribution of binaural response type and those related to integrated excitatory bandwidth. The more broadly tuned dorsal subregion of AI contained the largest number of EE neurons and the smallest number of EI neurons (61% vs. 35%, respectively). In the sharply tuned central region the percentage of EE and EI units was equal (46% vs. 47%). EI responses outnumber EE responses in the ventral subregion by approximately 51% to 40%. EO units were more often encountered in the central and ventral subregions than in the dorsal portion. However, due to the small number of EO responses encountered in the three cases under consideration EO responses were omitted from the regional analysis.

For all binaural response types combined, the Q_{10dB} and Q_{40dB} values for the dorsal and ventral subregions of AI were significantly below those for units located in the central region, as can be seen in Table IV. The same subregional differences for Q_{10dB} and Q_{40dB} could also be seen for the EE units alone. While the EI units showed the same trend, statistically significant differences were only observed between the central and ventral subregions for Q_{10dB} and between the dorsal and central subregions for Q_{40dB}. Overall, the sharpness of tuning for the EI units appeared to be slightly higher than for the EE units, although statistical significance of this difference was only observed for Q_{40dB} in the ventral subregion. This finding of sharpness-of-tuning differences between the binaural classes within subregions of AI suggests that the difference seen for the global analysis (Table III) is not due to differences in spatial distribution between the binaural classes but rather reflects a general difference in their frequency resolution that is largely independent of spatial location.

Another response parameter that appeared to be highly correlated with binaural response type was broadband transient response. Tables III and IV indicate that the broadband transient response differed

17

significantly between EE and EI neurons. The global analysis (Table III) indicated that EE neurons responded more vigorously to broadband stimuli than did EI cells. This was confirmed in the regional analysis (Table IV) for the central and ventral subregions. However, in the dorsal subregion EI neurons appeared to respond better to broadband stimuli although the difference between EE and EI neurons did not reach statistical significance.

An analysis of interregional differences between the bandwidth-related measures appeared to indicate that the parameters encountered in the central region differed from those in the dorsal and/or ventral subregions for both EI and EE responses. Since the division of AI into three subregions was based on a bandwidth measure (see Figure 11), this finding is to be expected.

Tone Intensity Parameters (monotonicity, threshold, SRL, dynamic range). Across the entire mapped extent of AI, the monotonicity of EE, EI, and EO units revealed no statistically significant differences (Table III). However, EI units displayed a tendency to be more nonmonotonic than either EE or EO units (Table III). In the global three-case analysis, a significant difference between EE and EI thresholds was observed such that EI neurons were found to be more sensitive (i.e., had lower thresholds) than either EE or EO units (Table III). Significant differences were also seen for the SRL responses in which lower SRLs were obtained for EI units as compard to EE or EO units. With respect to dynamic range, EE neurons appeared to have a slightly broader range than those of EI neurons.

Regional analysis revealed that EI neurons were more nonmonotonic than EE neurons in the central and ventral subregions of AI. This difference was statistically significant in the ventral portion of AI and had a statistical trend ($p<0.1$) for the central region. A significant difference between the EE and EI unit thresholds was seen in the ventral portion of AI (Table IV) with EI neurons exhibiting lower thresholds than EE neurons. Differences between EI and EE neurons were most clearly demonstrable for the SRL. In this case, all three subregions showed significant differences, with the SRL of EI units being 7 to 13 dB below those of the EE neurons. While the dynamic range of EI neurons tended to be narrower for EI responses than for EE responses, the difference did not reach statistical significance in any of the three subregions.

A number of interregional differences between the intensity parameter responses was observed for EE and EI classes individually, as well as for both classes combined. For monotoncity and threshold, these differences indicate that unit responses located in the central region differed from those located in the dorsal and ventral subregions. For the SRL and the dynamic range measures, the dorsal subregion appeared to be different from the central and ventral subregions.

Temporal Parameters (FM parameters). The global analysis of temporal parameters showed only an EE-EI difference in the preferred FM speed with EI neurons preferring slightly slower speeds than EE neurons. In the regional analysis, this finding was confined to the central portion of AI, although the ventral subregion showed the same tendency. There were no significant differences between binaural response type for either latency or preferred FM direction.

Interregional differences were apparent for all temporal parameter responses. The onset latencies in the central region were the shortest, although this was only statistically significant in the four-case analysis. Preferred FM speed observed in the central subregion differed significantly from that in the ventral subregion for EE and EI cells and for both binaural response types combined. When both types are considered together, the dorsal subregion also appeared to differ from those in the central subregion. For EE units in isolation or combined with EI cells, the preferred FM direction of units in dorsal AI differed from those encountered in the central and ventral subregions of AI.

In summary, distinct differences in several response properties of EI and EE units were observed throughout the full extent of the isofrequency domain of AI as well as within subregions of AI. Generally, EI neurons tended to exhibit lower response thresholds, lower SRLs, a narrower dynamic range, a higher degree of nonmonotonicity, sharper tuning, weaker responses to broadband stimuli, and preferred slightly slower FM sweeps than EE neurons. This suggests that there are stronger monaural, *contralateral* inhibitory influences acting upon EI neurons than upon EE neurons.

Regional differences in monaural responses between the dorsal, central, and ventral portions of AI appear to be more pronounced for EE units than for EI units. This is reflected in the fact that 11 of 30 regional comparisons showed statistically significant differences for EE neurons whereas only 6 of 30 such comparisons yielded a statistically significant difference for EI neurons. For both binaural classes

19

combined, 15 of 30 comparisons revealed significant regional differences, confirming the view of systematic functional segregations in the isofrequency domain of AI.

Discussion

The purpose of the present paper was to determine and compare the spatial distribution in AI of onset latency and binaural response type and compare them to the 10 parameters previously described in this series of reports (Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993). Similar to what was observed in most other parameters, onset latency responses were distributed nonrandomly across the cortex such that cells with longer onset latencies were typically located in the dorsal and ventral aspects of AI while those with shorter latencies were more often found in the central region of the dorsal-ventral extent of AI. In concurrence with earlier reports (Imig and Adrian, 1977; Middlebrooks et al., 1980), the present study also showed that near threshold binaural responses appeared to be distributed as alternating elongated bands or patches of EE and EI cells. While the overall spatial distributions of onset latency and binaural responses were similar in most instances, the details of these spatial patterns varied considerably across cases.

Before discussing these results further, the issue of multi- versus single unit recording should be addressed. As this issue has been discussed in detail elsewhere (Schreiner and Mendelson, 1990; Schreiner et al. 1992; Mendelson et al 1993; Schreiner and Sutter, 1992; Sutter and Schreiner, 1995), it will be treated summarily here. Since multi-unit recording samples the electrical activity of several neurons within a small volume, it cannot be assumed that all neurons involved have identical response properties, nor can the potential contribution of thalamocortical afferents impinging on postsynaptic cortical cells be excluded (Schreiner and Cynader, 1984; Schreiner and Mendelson, 1990). Thus, both multi-unit onset latency and binaural response recorded at a given cortical site most likely reflect the dominant pattern of neural activity resulting from local integrated cortical processing at that location. Moreover, the systematic variation in multi-unit onset latencies and binaural responses observed along the dorsoventral extent of AI is more easily explained as the consequence of intrinsically distributed cortical processes rather than as the result of random variation in multi-unit responses.

20

In order to diminish the potential effects that single versus multi-unit recordings might impose on the results, comparisons of single and multi-unit onset latencies were made as a function of dorso-ventral location. In general, it appeared that the onset latencies were shorter for the multi-units than for single units. However, upon closer inspection, it was apparent that there was a greater scatter among the spatial distribution of onset latencies for single units than for multi-units. This increase in scatter of single unit responses versus multiple-unit responses has been observed in CF, Q_{10dB}, Q_{40dB}, threshold and monotonicity measures, particularly in the ventral region of AI (Schreiner and Sutter, 1992; Sutter and Schreiner, 1994). Despite the small difference in onset latency the overall spatial distribution for single and multi-unit onset latencies was similar.

Onset Latency

The average onset latencies recorded in the present study are characteristic of, and closely resemble those obtained in single unit studies (Phillips and Irvine, 1981; Phillips et al. 1985). As observed by these researchers, as stimulus intensity was increased, onset latency became shorter.

As mentioned above, a comparison between single and multi-unit onset latencies revealed that multi-units responded with shorter latencies. This is most likely due to the fact that a given unit cluster will consist of individual cells exhibiting both short and long latencies. Thus, when recording from a unit cluster, neurons displaying shorter latencies will bias the aggregate response and, thereby, cause the cluster to be assigned a correspondingly shorter latency value.

The present study also revealed that onset latencies were not randomly distributed in AI. In general, units with shorter onset latencies were located in the central region of AI while units with longer onset latencies were typically found in the more dorsal and ventral subregions of the cortex. The spatial transitions from shorter to longer latencies and vice versa were sometimes gradual (e.g., Fig. 4; #87-001) and other times quite steep (e.g., Fig. 5b; #87-518). In addition, there was some regional variability in onset latency values, particularly in the ventral portion of the mapped areas. Several investigators have recently reported a higher degree of variability of ventrally located AI units in response to other stimulus parameters (Reale and Kettner, 1986; Imig et al. 1990; Schreiner and Mendelson, 1990; Schreiner et al.,

1992; Mendelson et al., 1993). It has been suggested that cells located in the ventral portion of AI may be organizationally or functionally distinct from those located in the central and dorsal regions of the cortex (Reale and Kettner, 1986; Imig et al. 1990; Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993; Schreiner and Sutter, 1992). The greater variability in onset latency observed in the present study provides additional support for this suggestion. The spatial extent and location of aggregates with similar latencies appeared to vary from animal to animal as well as with position within the isofrequency domain of AI. This intercase variability was also observed in the spatial distribution of other response properties in this series of experiments (Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993).

The observation that onset latencies appeared to show a large scatter in the spatial distribution and varied as a function of dorsoventral location suggests that latency reflects mechanisms other than travel time differences along the basilar membrane. One possible contributing factor to longer onset latencies observed in the dorsal region of AI, is whether the unit has a multi- or single peaked response profile. Sutter and Schreiner (1991) found that in general, neurons with multi-peaked profiles respond with longer latencies than single-peaked neurons. These researchers also observed that the majority of multi-peaked neurons were located in the dorsal region of AI. Thus, the presence of multi-peaked neurons may help to account for the longer onset latency responses observed in the dorsal region of AI. The longer latencies that we observed in the ventral region are most likely due to some other neural mechanism such as an increase in scatter of CF (Schreiner and Sutter, 1992), stronger inhibitory influences, or inputs from areas with longer latencies.

The large local scatter of onset latencies found for multi- and single units in AI is similar to observations made for CF (Schreiner and Sutter, 1992), threshold, and nonmonotonicity (Sutter and Schreiner, 1994). This supports the hypothesis that the local processing networks in AI can draw information from widely distributed parameter ranges in the spectral (frequency. bandwidth and amplitude) and temporal (latency, FM) domain. The advantage of the richness of these locally computed parameters superimposed on more global gradients, may be related to the formation of transcortical neural assemblies involved in the distributed encoding of spectrally and temporally complex signals.

22

The orderly, global spatial distribution of onset latencies observed in the present study does not appear to be unique to area AI. Schreiner and Langner (1988) observed that onset latencies were systematically distributed in the central nucleus of the inferior colliculus (ICC) of the cat. However, one potentially important difference between the spatial distribution of onset latencies in these two auditory structures is that latency appears to be largely independent of CF in AI, at least over the CF range tested in this study, while in ICC the gradient of onset latencies appears to be only weakly correlated with CF, suggesting that the travelling wave mechanism may be contributing more to the latency in ICC than to the one observed in AI (Langner et al. 1987).

A systematic distribution of onset latencies has also been observed in other species such as the domestic chick. Heil and Scheich (1991) found that not only did onset latencies decrease with an increase in frequency, but that they were also topographically distributed within field L, the avian homologue of the auditory cortex.

Binaural Response Type

In the present study, alternating elongated patches or bands of EE and EI cells were distributed orthogonally to CF. This pattern of spatial distribution is consistent with observations reported by other investigators (Imig and Adrian, 1977; Imig and Brugge, 1978; Middlebrooks et al., 1980; Schreiner and Cynader, 1984; Kelly and Judge, 1994). More recently, Reale and Kettner (1986), found that in the ventral region of AI a somewhat different pattern of binaural responses emerged when tested with interaural intensity disparity (IID) stimuli as opposed to the more traditional binaural equal intensity stimuli. With IIDs, they found a disproportionate number of cells in the ventral subregion that exihibited both facilitatory and inhibitory response properties when tested with IIDs as compared to equal intensity stimuli. This segregation in binaural responses between the ventral subregion and the rest of the cortex lends further support to the notion that there may be a functional partition between cells located in the ventral portion of AI as compared to those located more centrally or dorsally.

Anatomical support for the segregation of binaural resopnse types can be derived from earlier reports which have shown that the thalamic projections to AI are topographically organized (Anderson et

23

al., 1980; Merzenich et al., 1982; Middlebrooks and Zook, 1983; Morel and Imig, 1987). Middlebrooks and Zook (1983) showed that different EI bands receive input from coincident thalamic populations while different EE bands derive input from different subdivisions of a continuous EE-projecting zone. However, Brandner and Redies (1990) have recently suggested that cortical EE and EI bands do not receive coincident thalamic projections. Instead, their data suggest that different subregions of a cortical isofrequency contour may receive inputs from different portions of the corresponding thalamic isofrequency contour. Brandner and Redies (1990) reported that the rostral part of the ventral nucleus of the medial geniculate body (MGB) projects to neurons located in the dorsal region of the cortex while the caudal region of the MGB projects to more ventrally located cortical cells. These geniculate projections appear to be independent of binaural response properties and may provide some insight into the functional gradients that have been observed in the isofrequency domain (Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993; Heil et al., 1992). However, a more thorough re-evaluation of the thalamocortical projection patterns is now required in light of the functional distinctions which are apparent between different subregions of AI.

Comparison of the Spatial Distribution of Response Properties in AI

It is well known that the cortex is tonotopically organized and that alternating patches of EE and EI cells are distributed orthogonally to CF (Rose and Woolsey, 1948; Merzenich et al., 1975). Recently, several investigators have reported that a number of different response parameters appear to be organized in an orderly manner along the isofrequency axis of AI. For example, there have been reports which suggest that a subset of sound localization response properties may be spatially distributed within AI (Middlebrooks and Pettigrew, 1981; Jenkins and Merzenich, 1984; Reale and Kettner, 1986; Imig et al, 1990; Rajan et al, 1990). In addition, the spatial distributions of other stimulus response parameters, such as integrated excitatory bandwidth, broadband transient response, FM sweep direction and speed selectivity, stimulus intensity, inhibitory sidebands, and multipeaked responses, have also been related to the isofrequency domain of AI (Schreiner and Mendelson, 1990; Sutter and Schreiner, 1991; Schreiner and Sutter, 1992; Heil et al. 1992; Schreiner et al., 1992; Mendelson et al. 1993; Shamma et al., 1993).

24

For the purposes of the present discussion only those spatial distributions obtained from the current experimental series that are significantly correlated with latency or binaural responses will be discussed. A global correlational analysis of all twelve response parameters studied will be provided in the final report in this series.

Onset latency versus spectral, intensity, and temporal response properties recorded at the single cortical locations

Spectral Bandwidth.Parameters. In the present study a significant correlation between onset latency and Q_{10dB} was found for two of the three cases. This finding suggests that narrowly tuned cells respond with shorter onset latencies whereas more broadly tuned cells (at least for broadly tuned cells located in the dorsal region of AI) may respond with longer latencies. This is consistent with results reported by Kelly et al (1986) who found that broadly tuned cells in ferret auditory cortex had minimum first spike latencies that were 5-10 ms longer than more narrowly tuned cells.

A similar finding was observed for these same two cases described in the preceding paragraph when onset latencies were correlated with Q_{40dB}. A third case (#87-518) for which these correlations could be analyzed (and which were found to be nonsignificant) contained two regions in which units responded with shorter onset latencies and three regions where units displayed longer onset latencies The presence of these additional regions of shorter and longer onset latencies may have contributed to the nonsignificant correlation found between these two response properties for this case. Despite the significant statistical correlations between onset latency and Q_{10dB} and Q_{40dB}, the functional relationship between the integrated excitatory bandwidth and onset latency is not immediately obvious.

Onset latency was also found to be significantly correlated with broadband transient responses for all three cases studied. When a comparison of the spatial distribution of these two response parameters was made (Fig. 6; see also Schreiner and Mendelson 1990, Fig. 15) it became clear that those cortical regions containing cells responding most vigorously to broadband transients coincided with those regions containing cells exhibiting longer onset latencies.

25

Tone Intensity Parameters. Onset latency was found to be significantly correlated with monotonicity for two of the three cases examined. This is not surprising given that monotonicity is significantly correlated with Q_{40dB}. As mentioned above, cells with high Q_{40dB} values were found to respond with shorter latencies than cells with lower Q_{40dB} values. This suggests that cells with high degrees of nonmonotonicity tend to have slightly *shorter* onset latencies. This observation stands in contrast to the results of Phillips et al. (1985) who found that nonmonotonic cells had significantly longer mean onset latencies than monotonic cells. One possible explanation for the apparent discrepancy between our results and those of Phillips et al. (1985) may be due to the fact that we adopted a different measure of monotonicity. We used a continuous scale for the assessment of monotonicity, namely the slope of the high level segment of the rate/level function, as compared to Phillips and colleagues who used a discrete two-category classification scheme. Consequently, our measure resulted in a a larger number of cells being classified as nonmonotonic than reported by Phillips and colleagues. Thus, the larger number of nonmonotonic cells may yield different results with regard to onset latency. In addition, although it is true that many nonmonotonic neurons also have high Q_{40dB} values, this relationship does not hold for all AI neurons. Particularly in the dorsal and ventral subregions, fairly broadly tuned nonmonotonic neurons could be encountered. Many of these more broadly tuned nonmonotonic units located in the dorsal region of AI may also have responded with shorter latencies. Additionally not every sharply tuned neuron was nonmonotonic, i.e., of two neurons with equally narrow bandwidth which differ only in monotoncity, the one with the higher nonmonotonicity likely has the longer latency. Therefore, it can be concluded that the overall U-shaped distribution of onset latencies is in register with the distribution of sharpness of tuning. In addition, this distribution is locally modulated by influences related to the monotonicity of the response properties with more nonmonotonic neurons having slightly longer onset latencies.

Temporal Parameters. The present results showed that there was a significant correlation between onset latency and preferred direction of FM sweep. For two of the three cases, cells preferring downward-directed sweeps tended to respond with shorter onset latencies. Since downward sweeps start at the basal, short latency end of the basilar membrane, it is conceivable that the observed correlation is a reflection of that circumstance. However, the functional implications of this finding are, at present, unclear and may

26

reflec* more general (and presently not well understood) tempora! processing mechanisms governing FM sweep direction selectivity in AI.

Binaural response type versus spectral, intensity, and temporal response properties obtained at the same cortical locations

Before discussing these results, a potential difficulty with relating binaural response type to other response parameters within each case must be addressed. In order to accurately assess the functional relationship between binaural responses and other response parameters, the preferred method of analysis would be to correlate a continuous measure of binaurality with the other continuously distributed response parameters. However, unlike the other response properties which result in a continuum of responses (e.g., for Q_{10-dB} where there is a gradual progression from narrowly tuned in the central regions to broadly tuned in the ventral and dorsal regions of the cortex), binaural responses as identified in this (and typically in other studies) study form into discrete clusters: a given unit at a given cortical site is either EE, EI, or EO. Although the correlation between a continuous variable and a discrete variable can provide certain insights into their relationship, we restricted ourselves here to a comparison of the parameter distributions for the three non-continuously distributed binaural response classes.

Several investigators (eg. Merzenich et al., 1982; Middlebrooks and Zook, 1983) have suggested that contiguous pairs of EE and EI bands may form functional units within AI. This argument suggests that pairs of EE/EI bands should be analyzed separately from other adjoining pairs with regard to the other 11 response properties examined in our experiments. However, there is not yet sufficient evidence to support the hypothesis that a contiguous pair of EE/EI cortical bands is indeed a functional unit. Adopting this assumption prematurely would also not conform well to this type of analysis because the absolute number of cells encountered within a pair of bands may be quite small for some animals, as was the case in the present study, and, therefore, may not be suited for the kind of global statistical analysis employed here. Instead, we opted for combining the binaural analysis for four extensively mapped cases in order to increase the N and thereby strengthen the analysis. However, since the global picture of binaural band shape and distribution is quite different from animal to animal, we did not analyze the data band-by-band

27

but rather across all of AI or separately for the dorsal, central, and ventral subregions of AI. AI was divided into the three subregions based on the distribution of sharpness of tuning of multi-unit responses. A similar division has previously been successfully employed (Sutter and Schreiner, 1991, 1995; Schreiner and Sutter, 1992). The subregional analysis is necessary in order to distinguish between global differences between EE and EI neurons which are based on differences in their spatial distribution relative to some empirically determined response gradient (e.g. sharpness of tuning) and on those differences that are also locally discernable and that may correspond to differences in cell morphology and connectivity.

Finally, the interpretation of the observed differences, or lack thereof, between the binaural response types is complicated by two factors. First, a single parameter can vary over a considerable range across the isofrequency domain of AI as demonstrated by the nonuniform spatial distribution for most of the investigated parameters. Consequently, a large variance due to global spatial variations may obscure local differences between binaural response types in regions with much smaller local variance of a given parameter. Second, the nonuniform spatial distribution of the response parameters may covary with spatial variations of binaural response classes, so that global differences between parameters are due to biases in the spatial distribution of the binaural responses. For example, the significantly larger values of Q_{40dB} for EI neurons across three of the cases (see Table III) may be due to the fact that more EI responses are located in the sharply tuned regions of AI and more EE responses in the less sharply tuned regions of AI, while differences between EE and EI neurons in either the sharply or the broadly tuned regions may actually be negligible. It is difficult to avoid any biases from the spatial distributions of the various parameters, since the location and identity of different binaural aggregates is not uniform from one animal to the next (see Figure 9). Therefore, we could not perform an analysis separately for different well defined binaural regions such as, for example, α, β, or γ EI bands (Imig and Brugge, 1978).

Spectral Bandwidth Parameters. In the global analysis we observed that Q_{40dB} values for EI cells were slightly higher than values for EE cells suggesting that EI cells are more sharply tuned than EE cells. A trend similar to that seen for Q_{40dB} was observed for Q_{10dB} values, although it did not reach statistical significance. In the regional analysis, the difference in sharpness of tuning between EI and EE was most clearly expressed in ventral AI. Interregional differences in accordance with previously described

28

differences between dorsal, central and ventral AI were seen for both Q_{10dB} and Q_{40dB} for EE as well as EI neurons.

Another parameter closely related to the breadth of the tuning curve profile is the response strength to broadband transient stimuli (see Schreiner and Mendelson, 1990). As was observed with Q_{40dB}, a clear difference in the response strength of the EI versus the EE units was apparent with EI units exhibiting weaker responses globally as well as ventrally and centrally. In contrast, EI units in dorsal AI exhibited stronger responses to broadband stimuli than did EE units.

Tone Intensity Parameters. All four parameters that assess the response of cortical neurons to tones of varying intensity levels revealed differences between EE and EI sites both in the global and the regional analysis. The most consistent differences were seen for the SRL, where the EE units showed relatively higher values in each region. A similar trend was seen for response thresholds, suggesting that EI cells may have lower contralateral monaural thresholds that are at least 5 dB lower than EE cells. Schreiner and Cynader (1984) reported a similar result in their investigation of multi-units in AI and AII. This relationship is perhaps not too surprising when one considers that both binaural responses and threshold responses of high and low threshold aggregates were scattered throughout the entire extent of the mapped area. The functional significance of this relationship awaits further investigation.

Monotonicity values for EI units located in the central and ventral subregions as well as across the entire cortical region examined, were generally lower than those of EE units. This suggests that EI cells are more nonmonotonic than EE cells. This is in contrast to Phillips et al (1985) who reported that while monotonic and nonmonotonic cells were segregated along the dorsoventral extent of AI, they did not appear to coincide with any particular binaural response type. In other words, they found that the spatial distribution of monotonicity appeared to be independent of binaural response. As mentioned above, we used a different method for assessing monotonicity from that used by Phillips and colleagues (1985), namely the slope of the high level segment of rate/level functions. Thus, our measure of monotonicity may be more sensitive to the finer differences that undoubtedly exist between binaural response types. Finally, the observed differences between EE and EI units with regard to monotonicity should be interpreted with some caution since they may strongly reflect cluster composition effects. Recently, Sutter and Schreiner

29

(1995) showed that the predictive value of multi-unit monotonicity measures for the underlying single unit properties is limited.

Temporal Parameters. With regard to FM sweep responses, binaural response types were correlated differently with preferred speed of FM sweeps. EE cells preferred faster FM sweeps than did EI cells. This relationship held in both the global analysis as well as for an analysis restricted to the central region of AI.

Collectively, seven of the twelve monaural response parameters investigated in this series of experiments displayed statistically significant differences between the two main binaural response types, EE and EI, either in the global and/or in the regional analysis. The response parameters that did *not* exhibit significant differences were Q_{10dB}, latency, preferred FM direction, and CF. Thus, it can be stated that EI units exhibited a) sharper tuning 40 dB above threshold, b) less responsiveness to broadband transient stimuli, c) lower minimum thresholds, d) lower SRLs, e) stronger nonmonotonicity, f) narrower dynamic range, and g) preferred slower FM speed than EE neurons. These results lead us to speculate that the functional distinctions between EE and EI neurons suggests the existence of two largely independent processing streams in the auditory cortex that may be distinguishable by the pattern of their thalamocortical inputs (Middlebrooks and Zook, 1983) and their callosal connections (Imig and Brugge, 1978). The EE stream appears more broadly tuned for spectral and intensity information and thus may emphasize integrative processing, while the EI stream seems more sharply tuned for spectral and intensity information, potentially reflecting greater differentiation of input signals.

Most response properties of EI neurons (including the fundamental basis of their nomenclature) seem to be compatible with the suggestion that these units have a slightly stronger inhibitory influence on the shaping of the overall unit response profile. It is unlikely, however, that this inhibition is provided through inadvertent (eg., through crosstalk) input originating from the ipsilateral inhibitory pathway. This is based on the fact that the acoustic crosstalk between the two ears with a closed sound delivery system is approximately -40dB: that is, the signal reaching the ipsilateral ear via crosstalk is at least 40 dB less than the contralateral signal. Although the crossed input signal may contribute to the response profile at higher intensities, it is unlikely that properties such as threshold itself, dynamic range, type of monotonicity, and

30

sharpness of tuning that have already been established at 20 to 40 dB above threshold are affected by this crosstalk. Thus, it is probable that these functional differences reflect genuine distinctions between the processing that occurs within these two projection pathways.

Finally, the present study has revealed what appears to be a difference in the global spatial distribution of the response parameters for the two binaural classes. While EE neurons expressed differences in the response parameters between the dorsal, central, and/or ventral subregions in 11 out of 30 possible comparisons (37%, see Table IV), this was only observed in 6 out of 30 cases (20%) for EI neurons. This observation suggests that the spatial differences for EE neurons are more strongly expressed than those for EI neurons. However, the relatively low number of EI neurons in our sample of dorsal AI may have contributed to a lack of significance in a number of possible comparisons. By using a larger sample size, more significant differences may become apparent for EI neurons as well.

As mentioned above, it has been argued (e.g., Merzenich et al., 1982; Middlebrooks and Zook, 1983) that binaural bands in the high-frequency domain reflect some kind of specialization of different populations of neurons with, for example, emphasis of coding of spatial information in EI locations and emphasis of other perceptually important aspects, such as spectral distribution, in EE bands. The global differences described here for the EE and EI sites support such a view. The question of why there are several EE and EI aggregates in AI may relate to the necessity of having EE- and EI-type processing carried out for different bandwidth, intensity and temporal conditions. By distributing several bands across the isofrequency domain, it is assured that regions of sharp and narrow tuning, regions of high and low thresholds or monotonicity and regions of rapidly changing frequency (eg. FM sweep speed) are equally covered.

Concluding Remarks

The present report provides a comprehensive description of the spatial distribution of the final two out of twelve stimulus response properties previously reported in cat AI (Schreiner and Mendelson, 1990; Schreiner et al., 1992; Mendelson et al., 1993). The results obtained from these experiments, as well as those from other investigators (Jenkins and Merzenich, 1984; Imig et al., 1990; Rajan et al. 1990; Heil et

31

al., 1992; Shamma et al., 1993) have provided us with a greater understanding of the functional organization of primary auditory cortex. In general, the majority of response properties observed in our experiments appear to be organized along the dorsoventral axis although this organization does not appear to be very strict (Schreiner and Sutter, 1992, Sutter and Schreiner 1995). This suggests that there may be local processing within a given isofrequency band for at least some of these properties. Support for this notion of local processing can be derived from the investigations into the intrinsic cortical connections of AI (Matsubara and Phillips, 1988; Luethke et al., 1989; Wallace et al., 1991). Matsubara and Phillips (1988) were the first to show that unlike callossal connections, labelled cells were found in a variety of binaural response types, regardless of the homogeneity of the injection site. They also found that labelled cells occurred in patches of the same or higher CFs. Wallace et al. (1991) found that while labelled cells were often located within an isofrequency band, they could also be found in anterior and posterior locations. These results suggest that local processing may occur within an isofrequency contour and that this processing is largely but not completely independent of binaural interaction response.

The presence of a number of response parameters that appear to be systematically organized within AI may provide a set of criteria by which the dorsoventral extent of AI can be delineated. With the exception of binaural response, changes in the distribution of other response features, e.g., onset latency, are gradual on a global scale, but can be quite variable on a local scale. However, they may not necessarily provide an unequivocal set of criteria that would allow for a clear demarcation of the borders to areas neighbouring dorsal and ventral AI.

On a functional level, the systematic distribution of these (and perhaps other as yet unknown) response parameters may reflect the representational basis for the detection and identification of specific features of sounds arising from the natural environment. Schreiner and Sutter (1992) have recently suggested that the region in AId (dorsal to the Q_{10dB} peak) may be better suited for integrative analysis of broadband stimuli by responding in a relatively undifferentiated manner to tones or spectral peaks of different frequencies. This is supported by the observation that cells in this subregion tend to be more broadly tuned and to have a high responsiveness to broadband stimuli such as clicks. In contrast, AIv may be better suited for differential analysis of the spectral properties of broadband stimuli since units there are

32

characterized by sharp tuning which is distributed over varying ranges of CFs and which do not seem to respond to broadband stimuli as well as cells in dorsal AI.

Acknowledgements

References

Andersen RA, Knight PL, Merzenich MM (1980)The thalamocortical and corticothalamic connections of AI, AII and the anterior auditory field (AAF) in the cat: evidence for two largely segregated systems of connections. *J. Comp. Neurol.* 194: 663-701.

Brandner S, Redies H (1990)The projection from the medial geniculate to field AI in cat: organization in the isofrequency dimension. *J. Neurosci.* 10: 50-61.

Heil P, Scheich H (1991) Functional organization of the avian auditory cortex analogue. II. Topographic distribution of latency. *Brain Res.* 539: 121-125.

Heil P, Rajan R, Irvine DFR (1992) Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II: Organization of response properties along the 'isofrequency' dimension. *Hear. Res.* 63: 135-156.

Imig TJ, Adrian HO (1977) Binaural columns in the primary field (AI) of auditory cortex. *Brain Res.* 138: 241-257

Imig TJ, Brugge JF (1978) Relationship between binaural interaction columns and commissural connections of the primary auditory field (AI) in the cat. *J. Comp. Neurol.* 182: 637-660.

Imig TJ, Irons WA, Samson FR (1990) Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J. Neurophysiol.* 63, 1448-1466.

Jenkins WM, Merzenich MM (1984) Role of cat primary auditory cortex for sound localization behavior. *J. Neurophysiol.* 52: 819-847.

Kelly JB, Judge PW (1994) Binaural organization of primary auditory cortex in the ferret (Mustela putorius). *J.Neurophysiol.* 71: 904-913.

Kelly JB, Judge PW, Phillips DP (1986) Representation of the cochlea in primary auditory cortex of the ferret (Mustela putorius). *Hearing Res.* 24: 111-115.

Langner G, Schreiner C, Merzenich MM (1987) Covariation of latency and temporal resolution in the inferior colliculus of the cat. *Hearing Res.* 31: 197-202.

Leuthke LE Krubitzer LA, Kaas JA (1989) Connections of primary auditory cortex in the New World monkey, Saguinas. *J. Comp. Neurol.* 285: 487-513.

Matsubara JA, Phillips DP (1988) Intracortical connections and their physiological correlates in the primary auditory cortex (AI) in the cat. *J. Comp. Neurol.* 268: 38-48.

Mendelson JR, Schreiner CE, Sutter ML, Grasse KL (1993) Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. *Exp.Brain Res.* 94: 65-87.

Merzenich MM, Knight PL, Roth GL (1975) Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38: 231-249

Merzenich MM, Colwell SA, Anderson, RA (1982) Thalamocortical and corticothalamic connections in the auditory system of cat. In *Cortical Sensory Organization. Vol. 3: Multiple Auditory Areas*, Woolsey CN, ed. pp. 43-57, Human Press, Clifton, New Jersey

35

Middlebrooks JC, Dykes RW, Merzenich MM (1980) Binaural response-specific bands in primary auditory cortex (AI) of the cat: Topographical organization orthogonal to isofrequency contours. *Brain Res.* 181: 31-48

Middlebrooks JC, Pettigrew JD (1981) Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J. Neurosci.* 1: 107-120.

Middlebrooks JC, Zook JM (1983) Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex *J Neurosci.* 3: 203-224.

Møller AR (1977) Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli. *J. Acoustic. Soc. Amer.* 62: 135-142.

Morel A, Imig TJ (1987) Thalamic projections to fields A, AI, P, and VP in the cat auditory cortex. *J. Comp. Neurol.* 265-119-144.

Phillips DP, Irvine DRF (1981) Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. *J. Neurophysiol.* 45: 48-58.

Phillips,DP, Orman SS, Musicant AD, Wilson GF (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. *Hearing Res.* 18: 73-86.

Rajan R, Aitkin LM, Irvine DRF (1990) Azimuthal sensitivity of neurons in primary auditory cortex of cats.II.Organization along frequency-band strips. *J. Neurophysiol.* 64: 888-902.

Reale RA, Imig TJ (1980) Tonotopic organization in cat auditory cortex. *J. Comp. Neurol.* 192: 265-292

Reale RA, Ketmer RE (1986) Topography of binaural organization in primary auditory cortex of the cat: Effects of changing interaural intensity. *J. Neurophysiol.* 56: 663-682

Schreiner CE, Cynader MS (1984) Basic functional organization of second auditory cortical field (AII) of the cat. *J. Neurophysiol.* 51: 1284-1305

Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: Distribution of integrated excitation. *J. Neurophysiol.* 64: 1442-1459

Schreiner CE, Mendelson JR, Sutter. ML (1992) Functional topography of cat primary auditory cortex: Representation of tone intensity. *Exp. Brain Res.* 92: 105-122.

Schreiner CE, Sutter ML (1992) Topography of excitatory bandwidth in cat primary auditory cortex:single-neuron versus multiple-neuron recordings. *J. Neurophysiol.* 68: 1487-1502.

Shamma SA, Fleshman JW, Wiser PR, Versnel H. (1993) Organization of response areas in ferret primary auditory cortex. *J. Neurophysiol.* 69: 367-383.

Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65: 1207-1226.

Sutter ML, Schreiner CE (1995): Topography of intensity parameters in cat primary auditory cortex: single-neuron versus multiple-neuron recordings *J. Neurophysiol.* (in press).

Wallace MN, Kitzes LM, Jones EG (1991) Intrinsic inter- and intralaminar connections and their relationship to the tonotopic map in cat primary auditory cortex. *Exp. Brain Res.* 86: 527-544.

Woolsey CN, Walzl EM (1942) Topical projection of nerve fibers from local regions of the cochlea in the cerebral cortex of the cat. *Bull. Johns Hopkins Hops.* 71: 315-344.

Figure Legends

Fig. 1A-C. The spatial distribution of CF, integrated excitatory bandwidth, expressed as Q_{10dB}, and threshold in AI for case #86-173. Two dimensions of the pseudo three-dimensional projection represent the anteroposterior (a,p) and dorsoventral (d,v) extent of the cortical surface. The arrowheads along the anteroposterior dimension indicate the approximate orientation of the isofrequency contours. **A** Spatial distribution of CF. The elevation of the surface of the plot corresponds to frequency. As illustrated, CF increased monotonically as the electrode was advanced in a rostral to caudal direction. The sudden increase of CF in the ventral portion of the map indicates the transition to the second auditory field (AII; Schreiner and Cynader, 1984). **B** The spatial distribution of Q_{10dB}. The elevation of the surface plot corresponds to the magnitude of Q_{10dB} across the mapped cortex. Low values of Q_{10dB} correspond to regions where units had relatively broad integrated excitatory bandwidths while high values correspond to a region where units had relatively narrow bandwidths. The scale bar for the Q_{10dB} axis is only accurate for the nearest corner of the plot because of the distortion in the projection. As can be seen, Q_{10dB} was systematically distributed across the cortical surface studied. **C** The spatial distribution of minimum response thresholds. Elevation corresponds to the magnitude of the threshold. Thresholds were typically lower in the central region of the map. In addition, an increase of threshold along the cochleotopic gradient toward higher CFs was also observed. All data depicted in the pseudo three-dimensional plots were subject to a smoothing factor of 0.9 resulting in a small amount of additional smoothing (see Schreiner and Mendelson, 1990).

Fig. 2A-D. Representative examples of multiple unit onset latency responses as a function of stimulus intensity level re minimum threshold (0 dB). In general, onset latencies were longer at lower intensities.

Fig. 3A-D. Scatter diagrams of onset latencies as a function of characteristic frequency (CF). In general, onset latency appeared to be independent of CF.

Fig. 4. Spatial distribution of onset latencies obtained at 30 - 40 dB above threshold. The *upper portion* depicts a pseudo three-dimensional projection of the spatial distribution of onset latencies in AI for case #87-001. As with the pseudo three-dimensional projection of integrated excitatory bandwidth, two dimensions of the projection represent the anteroposterior (a,p) and dorsoventral (d,v) extent of the cortical surface. The arrowheads along the anteroposterior dimension indicate the approximate orientation of the isofrequency contours. The elevation of the surface of the plot corresponds to the onset latency (expressed in ms) across the mapped cortex. In the *lower portion* of the figure, the numbers in the box represent the actual onset latency values that underlie the pseudo three-dimensional spatial reconstruction. As illustrated, onset latencies were systematically distributed across the cortical surface examined.

Fig. 5A-C. Additional examples of the spatial distribution of onset latencies. Spatial distribution of onset latency for A case #86-173, B case #87-518 and C case #87-706. In all three cases, there was a nonrandom spatial distribution of onset latencies. Conventions are identical to those in Fig. 4.

Fig. 6A,B. Comparison of the spatial distribution of (A) single unit (SU) and (B) multiple unit (MU) onset latencies (determined at 30 dB above threshold) along the dorsoventral axis of AI. These spatial distributions are depicted in relation to the map derived for integrated excitatory bandwidth calculated at 40 dB above threshold whereby the maximum in the BW40 distribution was assigned the reference value of 0 mm. The positions of recording points dorsal (negative values) and ventral (positive values) are shown relative to this 0 mm position. In general, while the overall pattern of the spatial distribution for single and multiple unit was similar onset latencies tended to be shorter for multiple units than for single units.

Fig. 7A,B. Regional comparison of the mean and variance of onset latencies 30dB above response threshold for (A) single units (SU) and (B) multiple units (MU).

40

Fig. 8A,B. Regional comparison of the minimum latency for (**A**) single unit (SU) and (**B**) multiple unit (MU) responses. Minimum latency corresponds to the stimulus condition that elicited the shortest latency independent of stimulus intensity.

Fig. 9A-E. Spatial distributions of binaural responses for the four cases studied. A Schematic representation of the mapped area on the right hemisphere of case #86-173. The rectangle in the inset approximates the mapped area. The filled circles in the enlarged rectangle represent EE cells, open circles represent EI cells and filled triangles indicate EO cells. **B** Contour plot of the data illustrated in **A** with the stippled areas representing EE or EO cells and the white areas represent locations of EI cells. The dashed line indicates the border between AI and AII as determined by the integrated excitatory bandwidth measurements. **C-E** Contour plots of binaural responses for the three remaining cases.

Fig. 10A,B. Comparison of the spatial distribution of onset latency and binaural responses for (**A**) case #87-518 and (**B**) case #87-001. The shaded areas for the binaural response plots represent EE cells while the shaded regions for the latency plots indicate locations where units had onset latencies longer than 10.6 ms. The dashed line represents the orientation of the isofrequency axis for each case. In general, there was no covariance of onset latency and binaural response.

Fig 11. Distribution of Q-10dB values along the dorso-ventral extent of AI for case #86-173. The transitions from the central, more sharply tuned region toward the dorsal and ventral, less sharply tuned regions, serve as boundaries between the dorsal, central and ventral subregions of the isofrequency domain.

Table 1: Distribution of the average minimum response latency for binaural interaction classes.

===

	EE	EI	EO
Case			
#86-173	14.1 ±2.7 (93)	14.5 ±3.1 (64)	14.8 ± 1.8 (35)
#87-001	12.2 ±2.2 (40)	11.6 ±1.4 (46)	
#87-518	10.9 ±1.5 (44)	11.2 ±1.6 (31)	
#87-706	12.4 ±2.3 (41)	11.3 ±1.7 (31)	

Values are means ±SD; number of cortical locations given in parentheses.

EE, locations with excitatory/excitatory binaural interaction; EI, locations with excitatory/inhibitory interaction; EO, locations with only monaurally evoked responses. Only one case had sufficient numbers of EO responses.

Table 2: Correlation of minimum onset latency with cortical response measures obtained at the same locations.

	Case	AI		AId[a]		AIv[b]	
		r	P	r	P	r	P
Q-10dB	#87-706	-0.38	0.0003	-0.49	0.0005		
	#87-001	-0.31	0.002	-0.39	0.002		
Q-40dB	#87-706					-0.31	0.05
	#87-001	-0.34	0.0008	-0.39	0.002		
Step	#87-518	-0.25	0.03				
	#87-706					-0.50	0.002
	#87-001	-0.24	0.02				
THR	#87-706					-0.30	0.05
MON	#87-518	-0.34	0.005				
	#87-706	-0.46	0.0001	-0.48	0.001	-0.45	0.006
SRL	#87-706	-0.42	0.0001	-0.57	0.0001	-0.35	0.03
	#87-001					-0.37	0.03
DR	#87-001					-0.42	0.02
DS	#87-518	-0.29	0.009	-0.34	0.02		
	#87-706	0.28	0.05	0.33	0.03	0.34	0.04
	#87-001	0.27	0.007	0.31	0.02		

Linear regression analysis; r = correlation coefficient; P = level of significance (F-test).

[a]AId represents recording location in the portion of AI dorsal to the most sharply tuned multiple unit responses (Schreiner and Mendelson 1990).

[b]AIv represents location ventral to the most sharply tuned region of AI.

43

THR, threshold; MON, degree of monotonicity; SRL, strongest response level; DR, dynamic range; DS, direction selective response to FM sweeps; Step, response strength to a frequency step from 0.2 kHz to 64 kHz, i.e. broad-band transient (Schreiner and Mendelson 1990).

44

Table 3: Parameter distributions for three binaural response classes (three animals).

==

	N	Q-10dB	Q-40dB	Step (%)
EE	129	4.87 ± 3.05	1.87 ± 1.02 ⌐	64.1 ± 70.1 ⌐
EI	116	5.19 ± 2.94	2.38 ± 1.72 ⌐	43.1 ± 65.0 ⌐
EO	15	4.5 ± 1.97	1.81 ± 0.81	61.2 ± 86.3

	N	THR (dB SPL)	SRL (dB SPL)	MON (%/dB)	DR (dB)
EE	129	20.5 ± 12.8 ⌐	44.9 ± 16.8 ⌐	0.24 ± 0.69	20.1 ± 7.7 ⌐
EI	116	15.5 ± 11.1 ⌐	35.0 ± 15.8 ⌐	0.39 ± 0.81	17.9 ± 6.9 ⌐
EO	15	23.0 ± 15.3 ⌐	48.5 ± 15.3 ⌐	0.26 ± 0.68	18.5 ± 6.6

	N	Latency (ms)	PS	DS
EE	129	11.8 ± 2.4	2.20 ± 0.24 ⌐	0.05 ± 0.17
EI	116	11.4 ± 2.4	2.09 ± 0.24 ⌐	0.04 ± 0.17
EO	15	10.9 ± 1.4	2.15 ± 0.32	0.01 ± 0.17

Brackets indicate statistically significant differences (ANOVA, $p < 0.05$) between binaural classes.

PS, preferred speed. For further explanations see Tables 1 and 2.

Table IV: Regional distribution of parameter values for two binaural response classes.

		N	Dorsal	N	Central	N	Ventral
Q-10dB	All	83	4.39 ± 2.65	95	5.70 ± 2.88	84	4.82 ± 3.17
	EE	51	4.19 ± 2.58	44	5.65 ± 3.06	34	4.86 ± 3.52
	EI	29	4.77 ± 2.73	45	5.94 ± 2.84	43	4.70 ± 3.10
Q-40dB	All	82	1.66 ± 0.91	91	2.57 ± 1.71	82	2.01 ± 1.20
	EE	50	1.61 ± 0.96	41	2.46 ± 1.11	33	1.54 ± 0.64
	EI	29	1.78 ± 0.87	45	2.75 ± 2.16	43	2.39 ± 1.55
Step	All	82	55 ± 75	93	57 ± 69	83	51 ± 64
	EE	50	44 ± 63	42	76 ± 75	34	78 ± 69
	EI	29	66 ± 85	45	42 ± 64	41	29 ± 42
THR	All	82	20.7 ± 12.6	95	15.1 ± 11.3	85	20.1 ± 12.3
	EE	50	21.9 ± 13.9	44	16.4 ± 11.9	35	23.9 ± 11.1
	EI	29	17.8 ± 10.1	45	13.1 ± 10.4	42	16.6 ± 12.3
SRL	All	77	47.3 ± 15.2	93	35.5 ± 16.3	80	39.7 ± 17.6
	EE	46	49.1 ± 14.8	43	39.1 ± 17.8	31	46.8 ± 16.4
	EI	28	41.9 ± 13.6	45	31.9 ± 14.5	42	33.6 ± 17.3
MON	All	76	-0.18 ± 0.74	88	-0.52 ± 0.74	71	-0.24 ± 0.77
	EE	44	-0.28 ± 0.61	39	-0.37 ± 0.72	25	-0.01 ± 0.73
	EI	29	-0.06 ± 0.9	44	-0.63 ± 0.76	39	-0.36 ± 0.71
DR	All	76	20.5 ± 6.2	89	18.4 ± 7.5	71	17.9 ± 8.1
	EE	44	21.0 ± 6.6	40	19.4 ± 8.1	25	19.8 ± 9.0
	EI	29	20.2 ± 5.4	44	17.7 ± 6.9	39	16.3 ± 7.6

Table IV (continued)

		N	Dorsal	N	Central	N	Ventral
Latency	All	81	11.8 ± 2.1	94	11.2 ± 2.4	84	11.6 ± 2.5
	EE	49	12.2 ± 2.3	43	11.4 ± 2.7	34	11.6 ± 1.9
	EI	29	11.3 ± 1.7	45	11.1 ± 2.2	43	11.7 ± 2.9
PS	All	83	2.22 ± 0.23	95	2.04 ± 0.28	84	2.21 ± 0.28
	EE	51	2.21 ± 0.24	44	2.14 ± 0.23	34	2.27 ± 0.24
	EI	29	2.20 ± 0.20	45	1.93 ± 0.30	42	2.18 ± 0.30
DS	All	83	0.01 ± 0.17	95	-0.09 ± 0.16	84	-0.04 ± 0.15
	EE	51	0.01 ± 0.17	44	-0.09 ± 0.16	34	-0.08 ± 0.15
	EI	29	-0.02 ± 0.17	45	-0.08 ± 0.17	42	-0.02 ± 0.15

Brackets indicate differences (ANOVA, $P<0.05$) between regional or binaural parameter distributions. For further explanation see Tables 1 and 2.
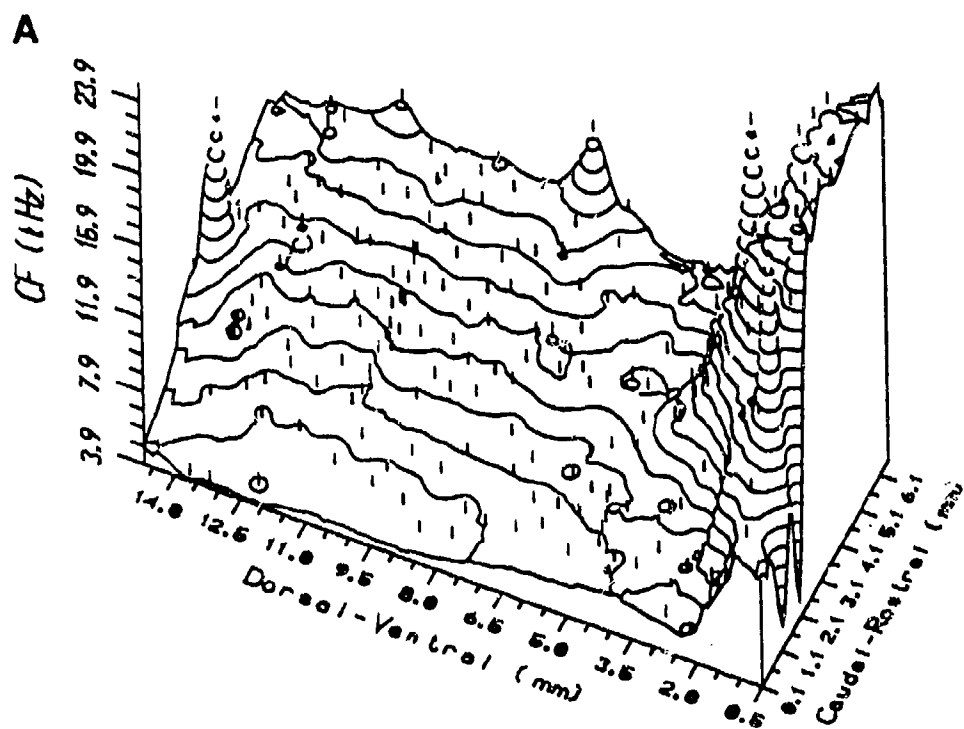
A

CF (kHz)
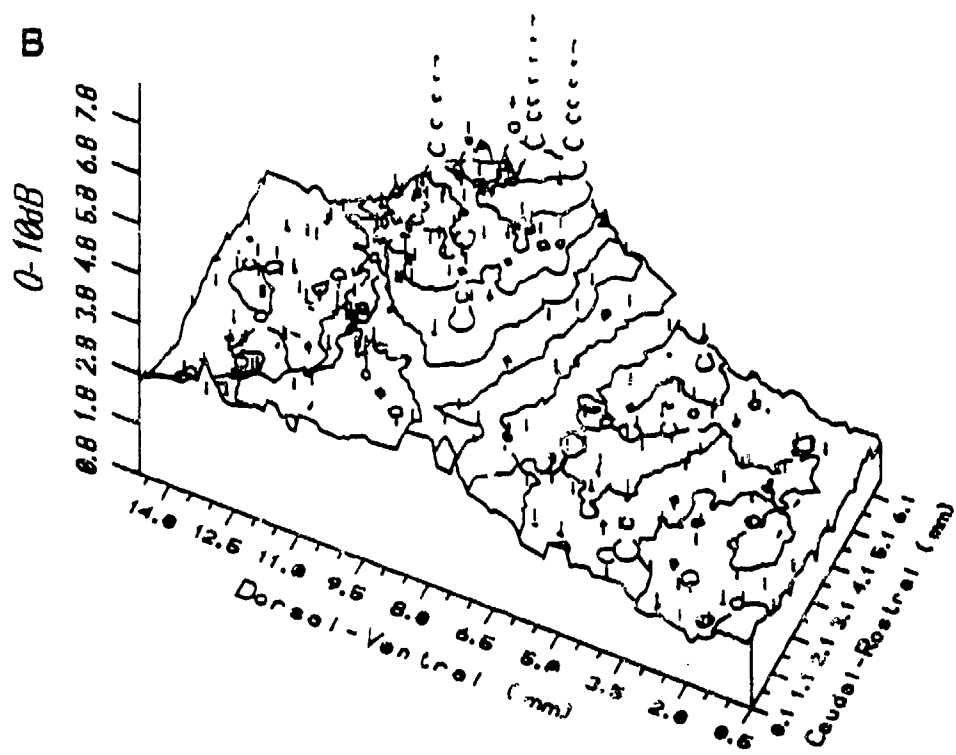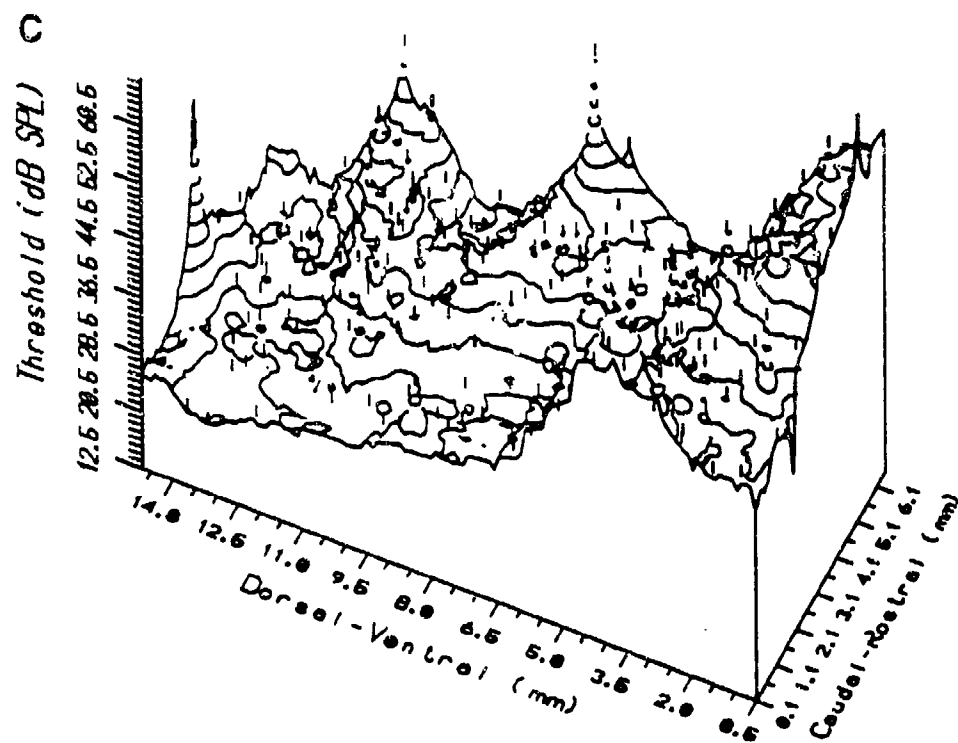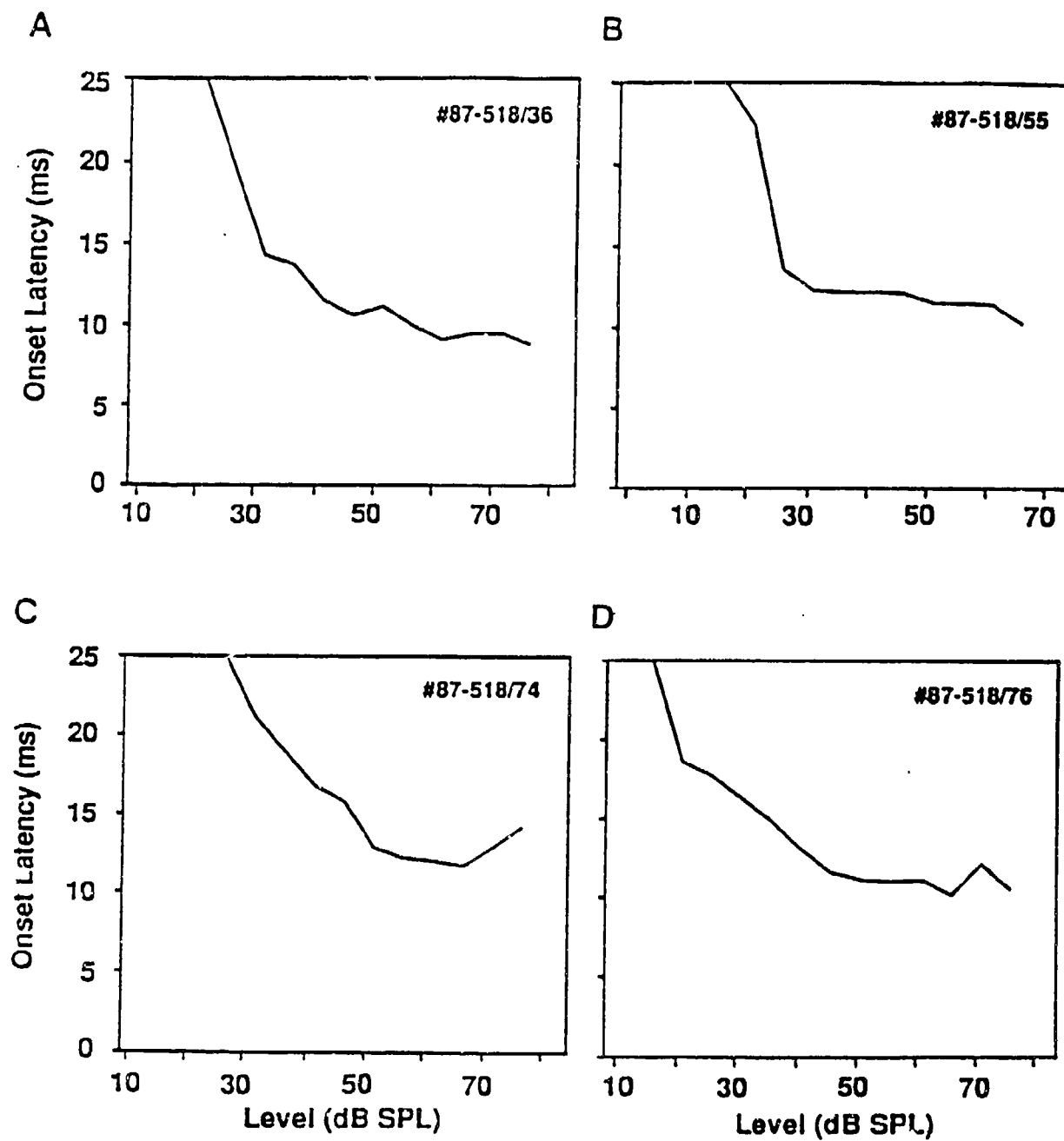
Fig.1A

B

0-10dB

Dorsal-Ventral (mm)

Caudal-Rostral (mm)

Fig.1B

C

Threshold (dB SPL)

12.5 20.5 28.5 36.5 44.5 52.5 60.5

Dorsal-Ventral (mm)

14.0 12.5 11.0 9.5 8.0 6.5 5.0 3.5 2.0 0.5

Caudal-Rostral (mm)

0.1 1.1 2.1 3.1 4.1 5.1 6.1

Fig.1C

Fig. 2

Fig. 3

#87-001

Fig. 4

A

Latency (ms)

#86-173

280 Lm

28 kHz

4 kHz

B

#87-518

Latency (ms)

480 µm

11 kHz

7 kHz

C

Latency (ms)

#87-706

450 µm

17.3 kHz

5.8 kHz

Fig. 5

Fig. 6

Fig. 7

A

SU

Minimum Latency (ms)

-3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3

Dorsal   Ventral

Location Relative to BW40 Map (mm)

B

MU

Minimum Latency (ms)

-3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3

Dorsal   Ventral

Location Relative to BW40 Map (mm)

Fig. 8

A

d

#86-173

500 μm

● EE
○ EI
▲ EO

v

p   a

B   #86-173

d

280 μm

A I

A II

v

p   a

EE

EI

EE

EI

EE

EI

C   #87-001

d

400 μm

A I

A II

v

p   a

EE

EI

EE

EI

EE

D   #87-518

d

460 μm

v

p AI AII a

EE

EI

EE

EI

EE

EI

E   #87-706

d

460 μm

A I

A II

p   a

EE

EI

EE

EI

EE

Fig. 9

#87-001

Latency < 10.6 ms

Binaural = EE

18 kHz

500 µm

#87-518

Latency
< 10.6 ms

Binaural = EE

10 kHz

500 µm

Fig. 10

Fig. 11

Spectral Envelope Coding in Cat Primary Auditory Cortex:
Properties of Ripple Transfer Functions (108)

by

Christoph E. Schreiner and Barbara M. Calhoun

Coleman Laboratory, W.M. Keck Center for Integrative Neuroscience, University of California at San Francisco, Box 0732, San Francisco CA 94143-0732.

Correspondence to: Christoph Schreiner, Coleman Laboratory, W.M. Keck Center for Integrative Neuroscience, University of California at San Francisco, Box 0732, San Francisco CA 94143-0732.
Telephone: (415) 476-2591; Telefax: (415) 476-1941; e-mail: Chris@phy.ucsf.edu

running headline: Spectral Envelope Coding in A1

ABSTRACT

Most of the functional aspects of the primary auditory cortex have been systematically explored using narrow-band stimuli. Naturally occurring sounds are usually spectrally complex and not easily characterized by narrow-band functions. In this study, the cortical response to a class of spectrally complex broad-band stimuli is systematically explored. The stimulus is a harmonic series of components whose spectral envelope is sinusoidally modulated. We show that responses of neurons in the primary auditory cortex are influenced by a number of spectral envelope parameters including the spacing of peaks in the spectral envelope, the phase of the envelope, and its modulation depth. Neuronal responses are also influenced by carrier properties such as the specific spacings of the harmonic components, the total bandwidth of the stimulus and the overall intensity. There is little apparent difference between the spectral envelope of transfer functions, phase response profiles, and intensity response profiles for single and multiple units. The results suggest an additional approach for characterizing responses of neurons to broad-band stimuli that are modeled after a behaviorally relevant class of sounds. Categorization of different types of responses to these stimuli is possible and can complement and expand the classical description of cortical receptive fields.

INTRODUCTION

The structure of naturally occurring sounds is frequently temporally and spectrally complex. Much of the exploration on the functional organization of the auditory cortex in mammals, however, has relied on the response to more simple stimuli such as pure tones (e.g., Merzenich et al. 1975; Reale and Imig, 1980; Phillips and Irvine, 1981; Phillips et al., 1985; Schreiner and Mendelson, 1990; Schreiner et al., 1992), and more parametrically accessible stimuli such as amplitude modulated tones or noise (e.g., Phillips and Hall, 1987; Schreiner and Urbas, 1988; Phillips et al., 1989; Eggermont, 1993), and frequency modulations (e.g., Suga and Jen, 1976; Mendelson and Cynader, 1985; Heil et al., 1992b; Mendelson and Grasse, 1992; Mendelson et al. 1993). Using these stimuli, a number of organizational principles of auditory cortical fields, particularly of the primary auditory field (AI), have emerged that are likely to provide the basis for the processing of the more complex stimuli in the acoustic biotope of each animal. Among the findings were the discovery of a systematic frequency representation in AI (e.g., Woolsey and Walzl, 1942; Merzenich et al., 1975), a non-uniform spatial distribution of the bandwidth of frequency tuning curves (Suga, 1965; Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992; Heil et al., 1992b), a non-uniform distribution of intensity related responses (Suga, 1977; Asanuma et al., 1983; Phillips et al., 1985; Schreiner et al. 1992; Heil et al., 1992b), a non-uniform distribution of inhibitory sideband characteristics (Shamma et al., 1992; Sutter et al., 1992), preferential responses to specific frequency sweeps (Shamma et al., 1992; Mendelson et al., 1993), regions responsive to different binaural interaction cues (Imig and Adrian, 1977; Middlebrooks et al., 1980) and regions sensitive to different spatial cues (Imig et al., 1990, Rajan et al., 1990). Each of these studies directly describes receptive field properties for coding specific stimuli features including spectral, temporal and spatial characteristics. Similar observations have been made in the equivalent area , Field L, of birds (Langner et al, 1981; Hose et al., 1987; Heil et al., 1992a).

However, the relationships between these basic functional organizations of cortical fields and the neuronal responses to natural signals, as well as the coding principles of complex signals, such as environmental sounds and communication sounds, have not been well established. It is not yet evident whether the receptive field properties described for the coding of simple spectral, temporal, and sound localization cues are suitable and sufficient to predict the neuronal response to a large variety of complex signals. This is largely due to three limitations. First, the description of cortical coding properties has concentrated individually on spectral, temporal and sound localization cues. For each of these cues, a number of descriptors

have been used to characterize the corresponding receptive field aspects. However, it is not clear which of the currently used descriptors are necessary for a complete characterization of the receptive field, and which of the parameters provide redundant or limited information. For example, the intensity dependence of tonal response is usually characterized by the response threshold, the dynamic range, the best level, and/or an estimate of the monotonicity of the rate/level function. However, even a combination of all four descriptors is not sufficient to fully reconstruct the rate/level function and make predictions of the firing rate for arbitrary soun' levels (see Raggio and Schreiner, 1994).

Second, interactions between the three main aspects of auditory receptive fields, i.e., spectral, temporal, and spatial, have been insufficiently addressed. In other words, it is not clear whether the spectral, temporal, and spatial coding mechanisms operate independently, and whether they can be characterized independently of each other.

And third, the high degree of temporal and spectral complexity of many natural or artificial signals renders the systematic definition and classification of relevant stimulus dimensions in those sounds difficult. Consequently, stimulus dimensions to completely and appropriately describe receptive fields in a systematic manner have not been fully identified.

This study begins to explore the cortical coding of a particular stimulus characteristic of broad-band stimuli – the spectral envelope. It uses several characteristics of the spectral envelope that potentially carry much of the spectral information including spacing between peak intensities, depth of modulation, overall intensity and bandwidth. Many auditory stimuli including most human vocalizations, music and other naturally occurring and artificial sounds, are characterized by broad-band spectra with distinctive, non-uniform spectral envelopes. Spectral envelopes are often essential for classification as demonstrated by the distinct perceptual differences between various vowel sounds. Vowel spectra are characterized by several frequency regions with increased spectral energy (formants) separated by regions with decreased spectral energy. Different vowels are formed by altering the formant spacing and/or position (e.g., Dickson and Maue-Dickson, 1982). Fig. 1 shows a spectral envelope for a human vowel sound and for a cat vocalization. Both vocalizations are marked by distinct maxima and minima in their spectral envelope. Recent evidence from studies on the sharpness of tuning (Schreiner and Mendelson, 1990, Sutter and Schreiner, 1991; Schreiner and Sutter, 1992) and the distribution of inhibitory sidebands (Shamma et al., 1993) suggest that cortical neurons are able to distinguish between different spectral shapes.

One approach to generalize and parameterize this stimulus characteristic is the use of broad-band spectra that have sinusoidal spectral envelopes with regard to logarithmic frequency and intensity scales. The benefits of 'ripple spectra' as exploratory stimuli lie a) in an easy characterization of the spectral envelope that is parametrically accessible and resembles features of naturally occurring complex signals, b) in the fact that the stimulus characterization is largely independent of the spectral content and the location of the 'carrier' signal, c) in its use as a tool for the exploration and characterization of the spectral receptive field using a broad-band stimulus, d) in its potential use to be able to compare these characterizations with those obtained with traditional methods such as single- and two-tone stimuli (Calhoun et al., 1993; Shamma et al., 1994) , and e) in its potential use as a stimulus to conduct a system-theoretical approach to the central auditory system that is equivalent to the spatial-frequency analysis used in the visual system and that may lead to an understanding for the processing of arbitrary spectral envelope wave forms (Hillier, 1991).

The spectral envelope of a ripple stimulus shares several important characteristics with luminance gratings commonly used to study processing in the visual system. Both stimuli provide similar distributed excitations across the receptor surface: in the visual system, the magnitude of the stimulus varies sinusoidally across the retina while in the auditory system, the magnitude varies sinusoidally along the basilar membrane. Application of this method in the visual system has lead to classifications of neurons by their response to sinusoidal gratings. Preferences to spatial characteristics of the stimulus, such as the grating frequency, modulation, and orientation, revealed important properties of central visual processing and provided a way of characterizing those properties.

In the first part of this study, we concentrate on characterizing the activity of cortical neurons and groups of neurons in the primary auditory cortex of the cat in response to ripple spectra for a range of ripple parameters such as spectral envelope modulation frequency, phase, and modulation depth. The relationships of ripple spectrum responses to properties of spectral receptive fields as obtained with pure tones and the system-theoretically based derivation of spectral receptive fields from ripple transfer functions will be shown in forthcoming papers.

METHODS
Surgery and Animal Preparation

Data were collected from 22 healthy adult cats. Federal and institutional guidelin~s were followed in the care and use of the cats. The cats were pre-anesthetized using a mixture of ketamine hydrochloride (10 mg/kg) and acepromazine maleate (0.25 mg/kg). They were given dexamethasone sodium phosphate (0.25 mg/kg/24 h) to control brain edema, and atropine sulfate (0.25 mg/12 h) to control mucous production. The femoral vein was exposed, and a venous cannulation was ~erformed. After an initial dose of sodium pentobarbital (to effect, ~30mg/kg), an a. lexic, hydrated state was maintained through constant infusion of an 8:1 mixture of lactated ringer's solution and sodium pentobarbital (~4 ml/h) with supplemental intravenous injections of sodium pentobarbital as needed. A tracheal cannula was inserted. A rectal temperature probe was used to record the temperature of the animal and maintain it at 37.5 C using a feedback controlled heated water blanket. The ECG and respiration rate was remotely monitored throughout the experiment.

The head of the cat was fixed with a mouth bar leaving the external meati unobstructed. The temporal muscle over the right hemisphere was then retracted. A cianiotomy was used to expose the lateral cortex above the ectosylvian sulcus. Finally, the primary auditory cortex was exposed by incising and reflecting the dura. The cortex was covered with silicon oil. For recording single units, the cortex was usually covered with a 1.5% solution of clear agarose in saline which diminished cortical pulsations while providing a fairly unobstructed view of the cortical surface.

Hollow ear bars were inserted into the ear canals to deliver the stimuli and a micromanipulator was positioned so that an electrode could be inserted perpendicular to the r ·rt · e of the cortex.

ElectrophysiClugy

Paryle : ~·/ea ~sten electrodes (Microprobe) with an impedance of 0.8 - 1.2 MOhms at 1 ! ¹lz wer inserted into the auditory cortex. A differential amplifier (World Precision Instruments DAM70-E) filtered the activity below 1 kHz and above 10 kHz. A window discriminator (BAK DIS-1) permitted acceptable threshold levels and wave form criteria to be set to distinguish action potentials from the background signal. In this manner, responses of a single unit, or of multiple units, could be selected and recorded. Activity of an acceptable size and shape resulted in a trigger pulse that was sent to a computer (DEC 11/73) that recorded the number and time of the pulses to predetermined stimuli for later analysis.

Acoustic Stimuli

Experiments were conducted in a double-walled sound-shielded room (IAC). Auditory stimuli were generated with a stand-alone digital signal processor (DSP TMS32010) with a 16-bit de-glitched DAC at a sampling rate of 60 or 120 kHz. The stimulus was low-pass filtered 96 dB/oct at 15 or 50 kHz, respectively. The amplitude of tonal stimuli was controlled by the number of amplitude steps in the generating wave form. Each step corresponded to ~0.15 mV resulting in a dynamic range of 70 dB. To ensure a good signal-to-noise ratio for the complex sounds, the full amplitude was used for those conditions. Passive attenuators (Hewlett Packard 350D) provided additional variable attenuation. The system was designed to provide a fairly flat transfer function when connected to the average cat ear (Sokolich 1981, US Patent 4251686). Headphones (STAX 54) were enclosed in small chambers and connected to sound delivery tubes. The tubes were inserted into the acoustic meati. Two different types of stimuli were used: tonal stimuli and ripple stimuli. The tonal stimuli were used to determine the frequency response areas, while the ripple stimuli were used to determine the transfer functions and response profiles to spectrally complex signals.

Frequency Response Area

Using a manually controlled frequency generator (General Radio; 1309-A Oscillator) to search for a response, neurons between 600μ and 1200μ were identified. Upon finding a responsive single neuron or neuron cluster, the first step was to determine the frequency response area (FRA). After making an initial estimation of the characteristic frequency (CF) and bandwidth of the receptive field by manually varying the frequency and intensity and using audiovisual response criteria, the FRA was obtained by pseudorandomly presenting stimuli from 15 intensity levels and 45 frequencies. The levels were steps of 5 dB, giving a sampled dynamic range of 70 dB. The frequency range was centered around the manually determined CF of the recording site, and covered between 3 and 5 octaves depending upon the estimated bandwidth. The frequencies were spaced in equal fractions of an octave over the entire range. Each stimulus was presented for 50 ms with a 3 ms rise time, and 400 ms inter stimulus interval.

Ripple Spectra

Once the FRA was determined, the stimulus was changed from pure tones to broad-band stimuli with distinct spectral envelopes. These stimuli will be referred to here as "ripple stimuli". The ripple stimulus is generated by sinusoidally modulating the spectral magnitude of a carrier along a logarithmic frequency scale. 100 to 256 harmonic signals were used as carriers. The fundamental frequency of these

harmonic components could be varied so that they usually ranged from 50 to 200 Hz and the maximum number of components was below 256 (the maximum number that could be produced by the DSP). The phase of the individual frequency components was fixed such that each component was phase shifted 53° farther than the previous one. This shifting eliminated strong peakiness in the temporal wave form that would have resulted with superposition of the components in cosine or sine phase. When plotting the spectrum on a double logarithmic scale, the spectral envelope of the stimulus was sinusoidal. Its bandwidth was set to 3 octaves with the geometric center of the band at a spectral maximum that usually was selected to correspond to the CF of the neuron. Since the individual components were linearly spaced, a 6 dB/octave decrease in the envelope of the ripple maintained a constant energy level per octave. The inverse wavelength of the sinusoidal spectral envelope is referred to as ripple density and is expressed in ripples/octave; the modulation of the spectral envelope is sinusoidal on a logarithmic scale and the standard spectral modulation depth (ripple depth) was 30 dB. The phase of the spectral envelope (ripple phase) is defined as 0° when the center peak of the spectral envelope at the geometric mean of the stimulus is aligned with the CF of the recording site. The overall intensity of the stimulus is expressed in dB SPL (as measured on the linear setting of a B & K sound level meter at the end of the ear bars). The stimulus had a duration of 100 ms including 5 ms rise and fall times. The interstimulus interval was usually 700 ms, and was increased to 1-1.5s if obvious adaptation effects were noted.

A schematic depiction of the standard stimulus, with a ripple density of 1 ripple/octave and a modulation depth of 30dB, is shown in Fig. 2A. Starting with the standard stimulus, individual parameters could be varied. For example, the ripple density can be changed (Fig. 2B) and the modulation depth can be varied (Fig. 2C). In addition, the overall intensity can be changed by increasing or decreasing each component by the same amount, and the fundamental frequency as well as the spectral width can be varied. As the ripple phase is varied, the central peak in the spectrum will move off the CF until a trough is aligned with the CF (Fig. 3), corresponding to a 180° phase shift. Note that envelope phase variations do not alter the frequency or phase properties of the carrier components, i.e., the location of the carrier spectrum is not shifted. The amount of frequency shift Δf (in octaves) of the apparent spectral peaks or troughs is determined by the ripple phase shift $\Delta\Phi$ and the ripple density (RD):

$$\Delta f = (1 / RD) * (-\Delta\Phi/360).$$

Ripple Transfer Function

Ripple transfer functions (RTFs) express the firing rate in response to a ripple spectrum as a function of ripple density. To find an appropriate intensity for the stimulus, a ripple stimulus was presented with a modulation depth of 30 dB and a ripple density of 1 ripple/octave and the overall intensity was varied until the level evoking the strongest response was determined using an audiovisual measure of response. This best intensity was used to obtain the ripple transfer functions.

Data were collected for an ordered sequence of ripple densities (15 values; 0 - 6 rip/oct; 0.2 to 1 rip/oct spacing) or a pseudorandom sequence (14 or 20 values; 0 - 8.66 rip/oct; 0.33 to 0.66 rip/oct spacing). Each stimulus condition was presented 25 times. In order to minimize adaptation to a particular stimulus, the inter stimulus interval was at least 750 ms, and occasionally over 1 second. In most cases, this long inter stimulus interval allowed the responses to the stimuli to be equally strong at the end of the 25 presentations as at the beginning.

The data were collected in post stimulus time histograms (PSTHs) allowing the number of action potentials and latency of each response to be determined. In order to create the ripple transfer functions from the PSTHs, the number of responses in each PSTH are determined. Since tonic responses were not seen, we only consider the number of spikes in a 30ms window that begins with the first spike after the onset of the stimulus.

Once the standard RTF was determined, response profiles were created for several other stimulus parameters. Response profiles plot the strength of the neuronal response as a function of a specific parameter and were obtained for the ripple modulation depth, the phase of the spectral envelope, and the overall intensity of the stimulus. In addition, effects on responses by the fundamental frequency and the bandwidth of the stimulus were investigated. In each case, PSTHs were obtained for 25 presentations of the stimulus. The PSTHs were analyzed and the response profiles plotted in the same manner as for the RTFs.

Analysis

From the RTFs and, if appropriate, the response profiles, the following properties were determined: the parameter values for the strongest or best response, the maximum and minimum spike response, the bandwidth of the response and the shape of the response. Based upon their shape, RTFs were categorized by the following filter types: band pass, low pass, high pass, notch and miscellaneous. To be classified, stimulus specific response modulations had to show at least a 30% difference from the best response. Intensity

response profiles were classified as monotonic or nonmonotonic according to the growth of the response at high intensities. Additionally, a modulation index of the RTFs and response profiles was determined by dividing the difference between the maximum and minimum response by the maximum response. This index varies between 0 and 1, and reflects the degree of response modulation due to changes in the varied parameter. The modulation index was determined for RTFs (Ripple Modulation Index, RMI) and for the phase profiles (Phase Modulation Index, PMI). The best response of a RTF or a response profile was the value of the variable parameter that elicited the strongest response. If two neighboring parameter values elicited identical responses, their values were averaged. The maximum response was the largest number of spikes that occurred following a stimulus onset. The minimum response was the average of the three parameter values that elicited the lowest number of spikes following stimulus onset. The bandwidth of a RTF is defined as the range of ripple densities over which the response is larger than the 50% value between maximum and minimum response.

Where possible, analyses were carried out individually for single unit and multiple unit recording sites and compared to each other. First, the variance and the mean of the two populations were determined. Then a variance ratio test was used to determine whether the variances of the two populations were the same, and a t-test was used to compare the means of the two populations. To increase the power of the test, and reduce the possibility of false negatives, $\alpha = 0.1$ was used for the test value. To compare the frequency distribution of filter classifications, a contingency table was set up and the chi-squared test used.

In addition to comparing single and multiple units, statistical analysis was used to compare the positive and negative phase symmetry indices by taking the absolute value of both indices, and comparing their variances and means. Again, the α-value was set to 0.1 to increase the power of the comparison.

RESULTS

Data were collected from 201 multiple units and 77 single units recorded in the primary auditory cortex (AI) of 22 adult cats. The first several multiple unit recordings were used to identify AI. Isofrequency contours identified the rostral and caudal boundaries of AI, while the sharpness of tuning (Schreiner and Mendelson, 1990) and threshold identified the ventral and dorsal regions (Schreiner and

Cynader, 1984; Schreiner et al., 1992). The recording sites spanned characteristic frequencies of 1 - 20 kHz, with the majority being in the 3 - 8 kHz range. Since most of the recordings were taken from the central two-thirds of AI, the rostral/caudal boundaries did not have to be precisely located. However, we did try to obtain responses from all regions of the isofrequency domain. The wide range of sharpness of tuning, with Q-10dB and Q-40dB values ranging from 0.5 to 10, indicates that a fairly diverse region, dorsal to ventral, was included in our sampling (Schreiner and Mendelson, 1990; Schreiner and Sutter, 1992).

Ripple Transfer Functions

Approximately 5% of the neurons identified in AI with pure tone stimuli did not respond to any of the ripple stimuli. All of these neurons were located in a narrow central region of AI, were extremely sharply tuned, and had highly nonmonotonic rate/level functions for pure tones. For each of the remaining 95% of the recording locations, a ripple transfer function (RTF) was obtained at a ripple phase of 0° and a spectral envelope modulation depth of 30dB. Figure 4A illustrates the construction of an RTF for a single neuron. The PSTHs for ten different ripple densities are shown. Each PSTH has been labeled with the ripple density of the stimulus and the number of action potentials elicited. For ease in comparison, the PSTHs have been placed in ascending order of ripple density and not in the pseudo random order in which they were presented to the animal. The ripple transfer function plots the number of spikes on the ordinate vs. the ripple density on the abscissa (Fig. 4B). The filled points represent one set of ten stimuli, the open points represent a second, complementary set of ten stimuli. The line represents a two-point average of the full data set. Several common properties of RTFs can be illustrated with this example including a best ripple density, the filter shape, the bandwidth of the response, and the modulation index. Commonly, the RTF reveals a range of ripple densities for which the response of the neuron is optimal. The best ripple density (BRD, the ripple density that produced the strongest response), is used as one of the descriptors of RTFs. In this case, the neurons have a best ripple density of 1 ripple/octave. The shape of the RTF in Fig. 4b can be classified as band pass since the responses to ripple densities below and above the best ripple densities are at least 30% smaller than at the best ripple density.

Another descriptor is found by determining the points below and above the peak where the response is half way between its maximum and minimum. The distance between these points, expressed in ripples/octave, is the bandwidth of the RTF at 50%. In this case, the bandwidth at 50% is 1.1 ripples/octave.

Finally, the ripple modulation index (RMI = (max. response - min. response) / max. response) was determined. In this example, RMI = (34 - 2)/(34) = 0.91. Figure 4c summarizes the various descriptors for a the RTF of a different neuron.

Figure 5 shows twelve representative ripple transfer functions constructed from 14-20 different ripple stimuli ranging from 0 to 8.6 ripples/octave. The transfer functions with two distinct symbols (Figures 5a, g, and i) are made of two pseudorandom sets of stimuli presented with a 15 second break between sets. The others are from stimuli presented in order from low to high ripple density. The six examples on the left (a-f) are typical RTFs showing band-pass filter characteristics. For these representative examples, the best ripple densities range from 0.7 to 5.0 ripples/octave. The six examples on the right (g-l) show representative examples of RTFs that are not band-pass filters. A notch filter is shown in Fig. 5g: ripple densities between 1 and 4 ripples/octave elicit almost no response while the responses to the highest ripple densities are within 50% of the maximum response. A low pass filter characterisitic is illustrated in Fig. 5h. In this particular case, the lowest ripple density presented was 0.3 ripples/octave. If lower ripple densities had been presented, there is a possibility this neuron would have turned out to be a band-pass filter. Double band-pass filters occur when there are two distinct peaks in the ripple transfer function with the secondary one at least 50% as high as the primary one (as seen in Figure 5l). For the purpose of this paper, there will not be a distinction made between double band-pass and band-pass aside from noting their existence. The RTF in Fig. 5j shows the characteristics of a band pass filter. Since the second peak is not within 50% of the first, it was not regarded as a distinct maximum. Finally, a high pass and an all-pass filter are shown in Fig. 5k and 5l, respectively. Obviously, in determining the type of filter, there is always some overlap, and approximation. For instance, unless a ripple density of 0 ripples/octave was presented, a low pass filter could actually be misclassified as a band-pass filter. If a filter did not have a clear classification, either because the data were too noisy, or because it overlapped with other categories it was considered 'miscellaneous'. All-pass filters (Fig. 5l) were also classified as miscellaneous. The majority of neurons (48%) were band-pass, while 18% were low pass, 6% notch filters, 3% high pass, and 25% 'miscellaneous' (Figure 6a).

No significant difference was found between the characteristics of RTFs of the single units (examples in Figs. 5b, c, d, e, g, h, and j) and multiple units (examples in Figs. 5a, f, i, k, and l). That there was little or no difference in the descriptors of single and multiple units can be seen in the summary histograms (Figure 6a: filter shapes; Figure 6b: best ripple densities). The majority of units, both single and

multiple, were band pass. Their best ripple density ranged from less than 1 ripple/octave to over 4 ripples/octave with a mean of 1.11 +/- 0.86 for single units and 1.25 +/- 1.16 for multiple units. To a significance level of $p = 0.10$, no difference was found for the two populations between the variances, the means, or the filtering characteristics (see Table 1).

The bandwidth of band pass RTFs can also be used to classify RTFs. A wide variety of bandwidths of RTFs was observed ranging from less than 1 ripple/octave (e.g., Fig. 5a) to over 3 ripples/octave (e.g., Figs. 5e and f). A frequency distribution for the bandwidth of the RTFs (Fig. 6c) shows that the bandwidth was usually around 0.5 - 1.0 ripples/octave although it was not unusual for the bandwidth to reach 3.0 ripples/octave. Again, the mean and variance for both single and multiple units were determined (see Table 1) and compared. To a significance level of $p = 0.10$, no significant difference was found.

The modulation index of the RTFs ranged from RMI = 1 (i.e., RTFs reflected a clear response to some ripple densities and no response to other ripple densities) to nearly zero (i.e., hardly any difference was seen between the responses to different ripple densities). Neurons that did not respond to ripple stimuli at all were not included. Figure 5c shows an example with a large modulation index (RMI = 1) whereas the RTF in Fig. 5l has a small modulation index of RMI = 0.36. Accordingly, there is a small difference between the maximum and minimum responses in the first case, and a large difference in the second case. The majority of neurons and neuron clusters had RTF modulation indices above 0.5. Mean values over all neurons were RMI = 0.81 +/- 0.16 for single units, and RMI = 0.71+/- 0.21 for multiple units (Fig. 6d), i.e., single unit RTFs showed a slightly higher modulation of response strength than multiple unit RTFs. The variances between the two populations, for $p = 0.1$, showed a significant difference indicating that the populations were different. However, there was no significant difference between the means (see Table 1).

In summary, the systematic investigation of the influence of a range of spectral envelope frequencies on cortical responses to broad-band stimuli revealed that a) the majority of cortical neurons are tuned to a specific range of spectral envelope frequencies; b) about half envelope transfer functions have band-pass characteristics; the remaining RTFs are low pass, high pass, notch, or of miscellaneous type; c) best ripple densities range from 0.3 to 5 ripples/octave with a mean around 1.2 ripples/octave; d) the sharpness of tuning to ripple densities varies between 0.3 and 3 ripples/octave with an average bandwidth of 1.3 ripples/octave; e) the degree of modulation of the response due to changes in ripple density of the stimulus can vary from nearly 0 to 100% with a mean of approximately 75%.

Phase Response Profiles

In the construction of RTFs, the stimulus was always set so that a spectral maximum, located at the geometric center of the stimulus, coincided with the CF of the neuron or neuron cluster under study. This phase constancy of the spectral envelope of the stimulus relative to the receptive field of the neuron(s) provides an important constraint for the interpretation of the described response dependences. It is apparent that other ripple phase choices, such as +180° or -180°, corresponding to a spectral minimum at the CF of the neuron, could result in quite different neuronal responses. In this section, the influence of the spectral envelope phase on the response to a fixed ripple density is evaluated. A phase shift of the spectral envelope corresponds to a frequency shift of the ripple maxima. A positive phase shift moves all peaks and troughs to lower frequencies while a negative phase shift will result in the movement of the peaks and troughs to higher frequencies. The frequency content of the carrier spectrum does not change with these phase shifts.

Figure 7 shows the phase response profile, the spike count vs. the ripple phase of the stimulus, for six different recording sites. In each case, a phase shift of 0° corresponds to a maximum aligned with the CF, and a phase shift of +180° or -180° corresponds with a minimum aligned with the CF of the recording site. The ripple densities that were used to determine the phase response profiles corresponded to either the best ripple density of the unit, or were approximately 1 ripple/octave since that was the average best ripple density. The actual ripple density used is stated in the legend. It is apparent that the best phase or the phase producing the strongest response in each example is at, or near, zero. Slight deviations from 0° of the maximum of several phase profiles can be partially attributed to small differences in the value of the CF at the threshold of the frequency response area and the best frequency at the actual intensity of the spectral peak.

In a linear system, the overall shape of the phase profiles would be cosine functions, indicated by the dashed lines in Figure 7. The observation that the phase profiles actually obtained are narrower than such functions and are also somewhat asymmetric reflects the operation of nonlinear elements in the processing of these stimuli. Three parameters will be used to characterize the shape of the phase profiles: the width at the 50% point, the symmetry of the profile at the 50% point, and the degree of modulation of the firing rate by phase shifts expressed as the phase modulation index (PMI, see Methods). The width of the phase profiles was commonly between 60° and 140° with a mean of 100° +/- 29°. This is significantly different from the width of a pure cosine function (180°).

For individual neurons, positive and negative phase shifts did not result in identical changes in the firing rate resulting in asymmetric phase profiles (see Fig. 10). This behavior is quantified by the Symmetry Index (SI). This index was found by determining the positive and negative phase shifts required to reduce the response by 50% and then calculating SI as follows: $SI = (\Phi_{pos} - \Phi_{neg})/(\Phi_{pos} + \Phi_{neg})$. Values of SI can vary from +1 to -1 with a value of zero reflecting a symmetric shape of the phase profile. Figure 8 shows the Symmetry Index for 11 single and 1 multiple units. The responses were nearly equally divided between positive and negative values. The average SI was -0.1 +/-0.4. By comparing the absolute values of the positive and negative asymmetries, we were able to determine whether the neurons, as a population, favored one direction, or behaved differently in one direction. The variance and the absolute value of the mean for the population of neurons with a positive asymmetry were not significantly different from those with a negative asymmetry (see Table 3).

Evaluation of the phase modulation index reveals that most phase profiles in our sample exhibited a full modulation of the firing rate by phase shifts. This modulation is similar to that seen for most RTFs. The average value for PMI was 0.77+/-0.22, and the mean and variance were not significantly different from those of the ripple modulation index.

In summary, the phase profiles revealed that there is a strong dependence of the responses on the position of the spectral envelope relative to the CF of the neuron. The relatively narrow width of the profile and its asymmetry strongly suggest the influence of nonlinear mechanisms such as compressive nonlinearities and nonuniformly distributed inhibitory influences that contribute to the processing of spectral envelope information.

Intensity Response Profiles

To determine RTFs and phase profiles, audiovisual criteria were used to set the intensity to produce a strong response. Usually, this intensity was approximately 15 to 25dB above the response threshold for the tested ripple density (ripple phase = 0°). To test the overall influence of the stimulus intensity on the response to ripple stimuli, rate-level functions were obtained for ripple densities at or near the best ripple density of the neuron or neuron group. Figure 9 shows six examples of rate-level functions. Neurons were classified as monotonic, nonmonotonic, or inconclusive. For nonmonotonic neurons, the firing rate peaked for a specific "best intensity", and then proceeded to decrease. Nonmonotonic neurons are defined as decreasing in their firing rate at least 30% with a 15 dB increase in intensity. Since

different neurons had different best intensities the intensity for some stimuli were never increased to levels large enough to determine monotonicity. Therefore, in addition to a definite decrease in spike rate, the greatest intensity had to be at least 15 dB greater than the spike rate that elicited the best response. Neurons which did not decrease their spike rate to below 70% for any intensity and did have at least a 15 dB difference between the best stimulus intensity and the greatest stimulus intensity were considered monotonic. Neurons that did not fulfill either requirement were classified as inconclusive (i.e. Fig 9a). In four out of the six examples (Fig. 9b, c, e, and f), the firing rate for the highest tested intensity was less than 70% of the maximum firing rate, i.e., they were nonmonotonic. The neuron shown in Figure 9d is clearly monotonic. The firing rate stays above 70% of the maximum firing rate for at least a 30 dB increase in intensity beyond the best intensity. Of the 42 units for which intensity response profiles were collected, only eight had clearly monotonic responses while twenty-five were nonmonotonic, and nine were inconclusive. Therefore, of the units where an identification could be made, 24% (8 out of 33) were monotonic, and the remaining 76% (25 out of 33) were nonmonotonic. Most of the units also had a narrow dynamic range, going from threshold to a peak response in 20 dB.

Modulation Depth Response Profiles

One of the principal descriptors of the spectral envelope is the depth of spectral modulation. In the ripple stimuli used to find the phase and intensity profiles, and the RTFs, the modulation depth was uniformly set to 30dB. To determine the modulation depth response profiles, the ripple density, intensity, and phase were set, and the modulation depth was varied. The ripple density was set at, or near, the best ripple density, the phase was set to 0°, and the intensity was the same as for determining the RTF and phase response. Then six to ten different modulation depths were presented between 0 and 40 dB. An increase of the modulation depth corresponded to a lowering of the troughs while keeping the levels of the spectral peaks constant. As a consequence, an increase in modulation depth resulted in a slight decrease in total energy (see Methods). Figure 11 shows six examples of modulation depth profiles. As shown in the first five examples (Fig. 11a, b, c, d, and e), the response usually increases as the modulation depth increases until a maximum or plateau is reached at depths ranging from 5 to 30 dB. As the modulation depth increased further, the firing rate was fairly independent of modulation depths. Figure 11f shows two modulation depth response profiles taken for the same unit. The closed circles were taken using the best ripple density of 1.0 ripple/octave while

the open circles used a ripple density of 3.0 ripples/octave. The clear differences between the two profiles indicate that the depth profile depends on the ripple density used. Figure 12 shows the frequency distribution for the lowest of the near optimal (best) modulation depths, at the best ripple density of the unit. The mean best modulation depth was 20+/-8dB. For most units, however, a modulation depth of 30dB still produced near optimal responses. Therefore, this modulation depth was selected for obtaining most RTFs and other response profiles.

Fundamental Frequency Response Profile

The ripple stimulus used in this study consisted of a series of harmonically related components. The fundamental frequency of the harmonic complex was chosen so that the complex contained at least 12ɔ components and maximally 256 components. As a consequence, the lowest octave of the 3-octave wide stimulus contained at least 18 and maximally 36 components thereby providing an adequate representation of the sinusoidal spectral envelope (See Figure 1). The fundamental frequencies actually used for obtaining RTFs and response profiles ranged from 37.5Hz to 150Hz. To systematically estimate the influence of variations of the fundamental frequency on the response to a ripple stimulus, we varied the fundamental frequency over a wide range for 12 neurons. The fundamental frequency response profiles in Fig. 13 illustrate that there were definite changes in the responsiveness as the fundamental frequency varied. For higher fundamental frequencies, the spectral envelope slope is under-sampled especially for high ripple densities and is no longer well represented. Therefore, it is reasonable to expect that at high fundamental frequencies, there are fairly large variations in the response magnitude. The reason for the variations at low fundamental frequencies (see Fig 13a, c, and f) is less clear, and suggests a reliance of the response on the fundamental frequency itself, rather than exclusively on aspects of the spectral envelope. Fundamentals ranging from 100 Hz to 1000 Hz (at which frequency the spectral envelope is being grossly under sampled) elicited strongest responses. No common pattern for fundamental frequency response profiles was apparent.

Spectral Width Response Profiles

Finally, the influence of the spectral width of the ripple stimulus was investigated. In all previous conditions, it was set to three octaves. This width was chosen to ensure that the stimulus covered the entire receptive field of the recording site. To evaluate the possible influence of the stimulus bandwidth on the response to ripple spectra, the bandwidth was changed systematically. Figure 14 shows six typical

spectral width response profiles. Two types of spectral width response profiles ca
be distinguished: those that generally decrease with increases in bandwidth (Fig
14a-e), and those that generally increase (Fig. 14f). Of the eleven units from which
we recorded the spectral width response profile, two (18%) were increasing, and
nine (82%) were decreasing. As the spectral width of the stimulus increases, the
response fairly monotonically varies until it reaches a steady state around a spectral
width of two to three octaves, i.e., wider bands do not significantly influence the
response. Therefore, the use of a three octave wide stimulus throughout the study is
a safe estimate of a bandwidth that covers most of the affected frequency receptive
field.

DISCUSSION

This series of experiments was conducted to investigate the response of
primary auditory cortical neurons to broad-band sounds with distinct spectral
envelopes. In particular, it was designed to determine whether neurons in AI are
sensitive to specific attributes of sinusoidal spectral envelopes, and which are the
most salient features of spectral envelopes that influence the cortical response.

We found that neurons in the primary auditory cortex can be tuned to specific
features of the sinusoidal spectral envelope of a complex sound including the spacing
of the spectral maxima and minima and the depth of spectral modulation or intensity
contrast. In addition, the position of the spectral envelope relative to the standard
receptive field is relevant, i.e., the intensity and the phase of the spectral envelope
affect the response. Carrier signal parameters of the spectral envelope signal, such
as bandwidth and spectral density, play a role in shaping the cortical response as
well.

Single-unit versus multiple-unit recordings

Before discussing the results of this study in the context of auditory coding
strategies, a brief consideration will be given to some methodological issues. Since
these are among the first parameteric experiments to explore the responses of
cortical cells to characteristics of the spectral envelope, we chose to record from both
single and multiple neurons. The main reason was the relatively large number of
stimulus parameters that needed to be systematically tested for each recording site.
Since the isolation of single units is more difficult to maintain over an extended
period of time, the use of multiple unit recordings yielded a higher percentage of
locations that were completely characterized for single tone, two-tone, and ripple
stimuli. The mean and variance of the best ripple density of recording sites, the

shape and bandwidth of the ripple transfer function, and the monotonicity of the intensity response profile were found to be quite similar for single and multiple unit recordings. Consequently, multiple and single unit results are discussed conjointly. However, the distinct differences between the two recording methods and their potential influence on the interpretation should be kept in mind (for discussion see Schreiner and Mendelson, 1990). Since some characteristics of pure tone frequency response areas, such as sharpness of tuning (Schreiner and Sutter, 1992) and monotonicity of rate/level functions (Sutter and Schreiner, in press) can show distinct differences between the results from these two recording methods, the similarity of the results obtained from single and multiple units for these broad-band stimuli is noteworthy. The similarity between the group response and the element response may be an indication that more complex stimuli, such as ripple stimuli, are better suited than pure tones for the study of local cortical properties, since they may engage and reveal more fully the cooperation within the local neuronal population. Since the responses were obtained from the middle laminae, no clear test of a potential columnar organization was possible. However, an indication of nonuniform spatial distributions along the isofrequency domain was evident suggesting that the spatial frequency analysis is locally and globally not randomly distributed.

Acoustic stimulus: carrier and spectral envelope

As pointed out in the Introduction, arguments for the use of ripple spectra in the exploration of cortical processing come from the study of the visual cortex, from recent psychoacoustical studies that focused on the coding of spectral envelopes, and from system-theoretical considerations.

Both temporally and spatially complex stimuli have been used for the highly advanced investigation of signal coding in the visual cortex (e.g., see Maffei and Fiorentini, 1973; Albrecht and DeValois, 1981). It has been shown in the visual system that certain tuning properties of a neuron vary with some other stimulus properties. The sharpness of directional orientation tuning, for example, is dependent upon whether the stimulus is a long bar or a wide bar. It has also been suggested that dividing neurons into categories based on response properties is dependent upon the stimulus used (Maffei and Fiorentini, 1976; Hammond and Munden, 1990). Of particular interest in this context are studies that utilize luminance gratings to explore the properties of visual cortical neurons (e.g., Maffei and Fiorentini, 1976; Zhang, 1990; Jagadeesh et al., 1993, for review see DeValois and De Valois, 1990). By using the response to sinusoidal lumin... :e gratings as a basis for a general

characterization, the response of many simple cells to various visual stimuli has proven to be predictable (Campbell and Robson, 1968; Worgotter and Eysel, 1987; Jagadeesh et al., 1993; De Angelis et al., 1993). In this case, stimuli excite a wide range of the receptor surface (the retina) in contrast to spatially restricted light points or bars. By manipulating the characteristics of the gratings, a number of receptive field properties can be evaluated that take into account long range and short range spatial influences and interactions. This approach is based on the system-theoretical equivalence of a system impulse response and system transfer or filter function applied to the spatial domain of the receptor surface. By presenting different spatial frequency gratings and taking the fourier transform of the resulting transfer function, estimations of the impulse response (the receptive fields) for simple cells can be obtained. Although investigations on the visual cortex have had great success determining and explaining receptive field properties by rigorously applying parametrically accessible stimuli under system-theoretical considerations, little of that approach has been transferred to the exploration of the central auditory system.

Early physiological studies investigating affects of the spectral envelope on auditory responses used a ripple-like stimuli, "cosine noise" (de Boer, 1967; Evans et al. 1970; Bilsen and Goldstein, 1974; Bilsen et al., 1975). By delaying white noise and adding it to itself, a broad-band stimulus was created with a sinusoidal spectral envelope, and a carrier of noise. The resulting stimulus had linearly spaced peaks and a spectral envelope amplitude that varied linearly. Although ripple-like, cosine noise is not shift invariant along either the frequency or the intensity axes and, therefore, is less suitable for our approach.

Later, psychophysical studies used spectrally complex stimuli to address the perception of spectral profiles (e.g., Bernstein and Green, 1988; Berg and Green, 1990; Green and Berg, 1991) and the influence of spectrally more distant regions on the processing and perception of presumably locally derived attributes (e.g., Hall, et al, 1984; Moore and Shailer, 1991; Richards and Heller, 1991). The sinusoidal spectral envelope has been used in a few psychoacoustical studies either to study the effects of spectral peak spacings (Houtgast,1977; Hillier, 1991; Keeling et al., 1993; Shamma et al., 1994), and to study spectral motion after-effects (Shu et al. 1993). Testing the effectiveness of a ripple stimulus as a forward or direct masker, Houtgast (1977) showed psychoacoustically that not only the spacing of the peaks, but also the phase of the spectral envelope is important. Investigating whether there are segregated "channels" in the auditory system for taking the fourier transform of the spectral envelope, Hillier (1993) found evidence for independent encoding of different ripple densities. While these studies reveal several similarities between the behavior of the

auditory system, and that of the visual system, when using comparable stimulus features, no directly corresponding physiological studies have been undertaken that could account for the psychophysical observations.

The success in the visual system of elucidating receptive field properties and perceptual phenomena with spatially complex stimuli that satisfy certain system-theoretical constraints, and the importance of spectral envelopes in discriminating vocalizations, particularly vowels, suggests a similar investigation of the spatial dimension of sounds, relative to the receptor surface, and their relationship to frequency organization. The results of such an approach provide an expansion of the possibilities to characterize neuronal response properties, should allow an assessment whether pure tone receptive fields are sufficient to predict the response to spectrally complex signals, and may reveal new insights into the principles underlying the representation of complex sounds in the auditory cortex.

Ripple Transfer Functions

Previous studies have used tonal stimuli to evaluate the receptive fields of neurons. Pure-tone receptive fields allow the characterization of the response selectivity of neurons by extracting descriptive parameters such as characteristic frequency, minimum threshold, and the bandwidth of the tonal response. Although such a characterization can be sufficient to predict the response of a peripheral neuron to a variety of stimuli, there is ample evidence that central neurons have stimulus selectivities that are not predictable from pure-tone responses alone. Examples include selective responses to complex sonar signals in bat auditory cortical neurons (Suga, 1965), vocalization specific neurons in squirrel monkeys (Newman and Wollberg, 1973a), song selective neurons in birds (Langner et al., 1981; Müller and Leppelsack, 1985), and syllable sensitivity in human cortical neurons (Creutzfeldt et al., 1989). In order to systematically explore potential response selectivities of cortical neurons beyond frequency and intensity, other stimulus dimensions that may be of relevance to the animal must be utilized. Transfer functions, intensity response profiles, and phase response profiles of ripple spectra provide a systematic characterization of neuronal responses that may reveal additional aspects of receptive fields while reflecting some of the properties seen with pure-tone responses. Although a thorough correlational analysis between RTF properties and pure-tone frequency response areas will be presented elsewhere, a brief discussion on the concepts and interpretations of RTF properties with regard to traditional tuning curves will be outlined. For the following discussion of observed RTF properties, it will suffice to consider only a few of the most common aspects of

frequency response areas, namely the width and strength of a central excitatory area and the position and strength of flanking inhibitory sidebands.

In this portion of the study, RTFs were exclusively obtained for the 'cosine' phase of the spectral envelope, i.e., a spectral maximum was located at the CF of the neurons. For this condition, four characteristics of RTFs were used to describe the variance in the obtained sample: the overall shape of the RTF, the ripple density producing the strongest response, the bandwidth of the RTF, and the degree of response modulation due to changes in the spectral envelope frequency.

The occurrence of the two most common types of RTF shape, band-pass and low-pass, representing 67% of the sample, can be largely attributed to the presence of relatively strong inhibitory sidebands in the FRAs. With the center peak of the ripple stimulus aligned with the CF of the unit, the maximum response to the stimulus will be attained when the neighboring spectral minima are positioned to coincide with the inhibitory sidebands. If the ripple density is lowered or increased, the invoked inhibition will increase by expanding the central peak into the sidebands or by moving neighboring ripple maxima into the inhibitory sidebands, respectively, thus creating a band-pass characteristic of the RTF. RTFs with low pass characteristics can be seen when inhibitory sidebands are absent or only weak. Other RTF filter shapes are probably also related to the spacing and relative strength of excitatory and inhibitory regions. Asymmetries in the arrangement and strength of these portions of the FRAs are likely contributing to the creation of notch or irregularly shaped RTFs. It has been shown (Suga and Manabe, 1982; Shamma et al., 1993; Sutter and Schreiner 1990), that the strength and relative position of the excitatory and inhibitory regions can vary substantially between neurons. Thus, the different receptive field organizations required to create the observed RTFs of varied shapes do exist.

Another important descriptor of RTFs relates to the position of the peak and/or the upper cut-off along the ripple density axis. These aspects of RTFs reflect both the width of the classical excitatory receptive field and the distance of the inhibitory sidebands from the center of the excitatory portion. Neurons showing a preference for higher ripple densities are likely to have narrow tuning curves and more closely spaced inhibitory sidebands. By contrast, preferences for low ripple densities probably reflect more broadly tuned neurons that, as a consequence, have fairly widely spaced inhibitory sidebands. The best ripple densities between 0.5 and 4 ripples per octave, as seen for the majority of neurons in this study, indicate that spacings between inhibitory and excitatory regions in neurons can vary between 0.25 to 2 octaves, consistent with findings for cortical neurons obtained in two-tone

experiments (Suga and Manabe, 1982; Shamma and Symmes, 1985; Sutter and Schreiner, 1990; Shamma et al., 1993).

Although it has been shown that the distribution of excitatory regions as well as that of inhibitory side bands and their relative strength can be asymmetrical, (Shamma and Symmes, 1985; Sutter and Schreiner, 1990; Schreiner and Sutter, 1992; Shamma et al., 1993), RTFs obtained by centering the ripple stimulus over the characteristic frequency of the neuron cannot necessarily resolve the direction of the asymmetry. However, by varying the phase of the spectral envelope for a given ripple density or by obtaining RTFs for different phases, some of these asymmetries can be investigated (see below).

The interpretive value of two other descriptors of RTFs, the bandwidth and the modulation index, is more difficult to assess. The modulation index is an indicator of the neuron's variation in activity due to changes in the ripple density of the spectral envelope. A high modulation index suggests a relatively high efficacy of excitatory or inhibitory influences of surrounding areas on the response strength of a neuron and/or a large change in the response strength with small (~ 6dB) changes in intensity. A low modulation index suggests relatively weak excitatory and inhibitory influences from surrounding areas, and/or not much change in the strength of the response with small changes in intensity. The bandwidth of RTFs also reflects these properties but, in addition, incorporates aspects related to the spacing of excitatory and inhibitory regions which are more directly expressed by the best ripple density or the upper cut-off density.

A useful feature to distinguish vowels independent of speaker gender or age is to consider the frequency ratio of the formants (e.g. Miller, 1989). Depending upon the vowel, most formants are spaced between 0.3 and 4.0 octaves with the majority of ratios cluster around 1 octave. The best ripple densities of the neurons in AI in this study covered a similar range with the majority tuned around 1 ripple/octave. Preliminary analysis of cat vocalizations (see Figure 1) also indicates, similar to human vowels, a preponderance of formant ratios around 1 octave.

Taken together, the characteristics of cosine RTFs reflect a number of receptive field properties that have been observed with pure-tone or two-tone measurements of FRAs. In addition, the RTF may provide information that is usually not reflected in FRAs, namely the cumulative effect of excitatory and inhibitory influences on the response strength including effects from regions outside the classical receptive field.

Response Profiles

The various parameters of the ripple stimulus that were controlled included the phase of the spectral envelope, the modulation depth, the spacing of the components, the bandwidth of the stimulus, and the overall intensity.

Variations of the phase of the spectral envelope invariably had a large effect on the magnitude of the response. It is known from pure tone stimuli that the characteristic frequency is bordered by regions that are either inhibitory or non-responsive. Thus, it would be expected that the strongest response to a ripple stimulus would be for a stimulus with a maximal amount of energy near the CF, and a minimal amount of energy in the bordering regions -- especially those corresponding to inhibitory areas. Since one of the goals of this particular stimulus was to cover the entire spectral range which might influence the response of a cell, the center peak of the stimulus was positioned at the CF of the unit, covering frequencies 1.5 octaves on either side. Changes in the ripple phase for a fixed ripple density should reveal whether this position was indeed producing the strongest response. Only very few phase profiles showed secondary peaks or a peak significantly shifted from the 0° position. That is, the alignment of a spectral maximum with the CF was near optimal. Slight deviations in the phase profile of the maximum from zero may be accounted for by discrepancies between the BF of a neuron at the stimulus intensity that corresponded to the local energy of the ripple stimulus and the near threshold estimate of the CF.

If the filtering system reflected in the RTFs were linear and symmetrical, the phase response profile for a given ripple density would be sinusoidal (the profiles would match the sinusoidal patterns -- hatched lines in Fig. 10). The clear deviations from the sinusoidal profile in all the phase profiles is a strong indication of either nonlinear components or asymmetries in the filter. Differences in the symmetry of the phase profiles are likely closely related to the asymmetries in the distribution and strength of inhibitory sidebands as shown with two-tone stimuli for AI neurons (Shamma et al., 1993; Sutter and Schreiner 1990, 1991). Equal numbers of positive and negative asymmetries suggest an equal number of neurons with stronger high and low frequency inhibitory sidebands (Shamma et al, 1993).

As the spectral modulation depth of a stimulus with fixed and near optimal ripple density was systematically increased, the responses usually increased up to a constant value. However, on some occasions responses decreased with increasing modulation depth. The most likely explanation for this difference in response behavior with changes in modulation depth is again related to the presence and strength of influence by inhibitory sidebands. If strong sidebands are present, an increase in the modulation depth of the optimally positioned spectral envelope will

mostly decrease the input to the inhibitory sidebands while maintaining the input to the central excitatory region. However, if no or only weak inhibition is present or if the spectral envelope is not optimally positioned, an increase in the modulation depth may remove excitatory energy from the receptive field and, thus, reduce the response strength. Most neurons showed the greatest changes in responsiveness for modulation depths between 5 and 25dB. For a reference on spectral modulation depth sizes, vowels usually have spectral modulation depths of 10 - 30 dB between formants, the same range where the neurons were most sensitive to change in the modulation depth of the spectral envelope.

As a function of overall intensity, the neurons predominantly behaved in a nonmonotonic manner in response to ripple stimuli. This behavior contrasts with that seen in response to pure tone stimuli. Using pure tones, various portions of sampled neurons in A1 (24-60%) have been reported to have shown nonmonotonic behavior (Phillips et al., 1985, 1989; Schreiner et al., 1992). The finding here, that 76% of the neurons showed nonmonotonic rate/level functions for ripple stimuli suggests that neurons are more selective or more sharply tuned to specific signal intensities when stimulated with broad-band stimuli than with narrow-band stimuli. Similarly, a high percentage of nonmonotonic rate/level functions have been found by Imig and colleagues (1990) for free-field wide-band noise stimuli. Both observations are consistent with the hypothesis that the activation of inhibitory sidebands plays a major role in creating intensity selectivity (Phillips, 1988).

As was seen, there are considerable changes in the response as the fundamental frequency of the carrier band varies over a limited range. As the fundamental frequency is varied, the number of components per octave or per critical band varies. However, as long as the fundamental frequency is small, the energy per octave or per critical band remains fairly constant. As the fundamental frequency increases, there are fewer components per octave or per critical band. In the extreme case, when there are only a few components per octave, the spectral envelope shape will not be adequately represented due to undersampling. At this point, the energy distribution over critical bands varies greatly with changes in the fundamental frequency. Since the local energy near the CF contributes significantly to the strength of the response, it might be expected that different high fundamental frequencies can produce large differences in the response. By contrast, different low fundamental frequencies should produce little, if any, variation in the response strength. Thus, the significant variations in the response strength, even between relatively low fundamental frequencies (50 to 100 Hz), are not consistent with such an energy distribution hypothesis. Alternatively, neuronal responses in

AI could be sensitive for specific values or value ranges of fundamental frequencies. The changes in the response to ripple spectra with fundamental frequency could, therefore, be due to changes in periodicity pitch that are encoded at the cortical level (Langner et al., 1994).

The final stimulus parameter that was systematically varied was the spectral bandwidth. Using a spectral width of 3 octaves as the standard bandwidth, we attempted to reach the entire spectral influence sphere of the neuron. Spectral regions far outside the pure-tone receptive field of the neuron were not expected to have significant effects on the neuron's response, and therefore, the stimuli used were band-limited. However, it was important to use a stimulus that was broad enough to cover the whole excitatory/inhibitory receptive field. The strongest responses were frequently seen at the smaller bandwidths and there was frequently a significant decline in the neuron's response as the spectral width of the neuron increased. Presumably, this occurred as the stimulus invaded the inhibitory regions of the receptive field. As the stimulus continued to spread, its boundaries were outside the receptive field, and further broadening did not affect the response of the neuron. For most neurons the bandwidth marking the transition to a steady response strength was approximately two to three octaves suggesting that influences from regions more than 1 - 1.5 octaves to either side of the CF are minimal.

The phase of the harmonic components was chosen to produce a nearly flat temporal envelope in order to avoid potential effects from peripheral nonlinearities observed for peaky wave forms (Horst et al., 1990). Although the potential influence of the component phase on the response to ripple stimuli remains to be investigated, preliminary observations indicate that the effects will be small compared to effects seen for variations of spectral envelope parameters.

In conclusion, broad-band stimuli with sinusoidal spectral envelopes were utilized for analyzing properties of auditory cortical neurons. The results show that the neuron's responses are sensitive, if not selective, to specific aspects of the spectral envelope such as the spatial frequency, ripple phase, or ripple depth. The results suggest that cortical neurons can represent the shape of broad-band stimuli in the form of a spatial fourier analysis of the spectral envelope. The conclusions drawn from the ripple transfer functions are limited by the use of only one ripple phase. In order to completely characterize a neuron's response, each spectral envelope frequency must be presented at least for two orthogonal phases (e.g. 0 and 90 degrees) so that it analyzes the symmetrical properties of the neurons response as well as the asymmetrical properties. In addition, any possible relationships between

traditional pure-tone tuning curves, and these response profiles must be investigated. We will be addressing these aspects in the future.

Bibliography

Abbs, J.H., and H.M. Sussman (1971) Neurophysiological Feature Detectors and Speech Perception: a Discussion of Theoretical Implications. *J. of Speech and Hearing Research* 14:23-36.

Aitkin, L. (1990) *The auditory cortex: Structural and functional bases of auditory perception.* Chapman and Hall, London.

Albrecht, D.G., and R.L. DeValois (1981) Striate cortex responses to periodic patterns with and without fundamental harmonics. *J. Physiology* 319:497-514.

Asanuma, A., D. Wong, and N.Suga (1983) Frequency and Amplitude Representations in the Anterior Primary Auditory Cortex of the Mustached Bat. *J. Neurophys.* 50(5):1182-96.

Berg, B.G., and D.M. Green (1990) Spectral weights in profile listening. *J. Acoust. Soc. Am.* 88(2)758-66).

Bernstein, L.R., and D.M. Green (1988) Detection of changes in spectral shape: uniform vs. non-uniform background spectra. *Hearing Research* 32:157-66.

Bilsen, F.A. (1977) Pitch of Noise Signals: Evidence for a "Central Spectrum". *J. Acoust. Soc. Am.* 61(1)150-61.

Bilsen, F.A. and J.L. Goldstein (1974) Pitch of Dichotically Delayed Noise and its Possible Spectral Basis. *J. Acoust. Soc. Am.* 55(2)292-6.

Bilsen, F.A., J.H. ten Kate, T.J.F. Buunen, and J. Raatgever (1975) Responses of Single Units in the Cochlear Nucleus of the Cat to Cosine Noise. *J. Acoust. Soc. Am.* 58(4):858-866.

Calhoun, B.M. and C.E. Schreiner (1993) Spatial frequency filters in cat auditory cortex. *Soc. Neurosci. Abstr.,* 23:581.8.

Campbell, F.W. and J.G. Robson (1968) Application of Fourier Analysis to the Visibility of Gratings. *J. Physiology* 197:551-66.

Creutzfeldt O., G. Ojemann, and E. Lettich (1989) Neuronal activity in the human lateral temporal lobe. I. Responses to speech. *Experimental Brain Research,* 77(3):451-75.

DeAngelis, G.C., I. Ohzawa, and R.D. Freeman (1993) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J Neurophys.* 69(4)1118-35.

de Boer, E. (1967) Correlation Studies Applied to the Frequency Resolution of the Cochlea. *J. Auditory Research* 7:209-17.

De Valois, R.L., and K.K. De Valois (1990) *Spatial Vision.* Oxford University Press, New York.

Dickson, D.P., and W. Maue-Dickson (1982) *Anatomical and Physiological Bases of Speech.* Little, Brown and Company, Boston.

Eggermont, J.J. (1993) Differential effects of age on click-rate and amplitude modulation-frequency coding in primary auditory cortex of the cat. *Hearing Research* 65(1-2)175-92.

Evans, E.F., J. Rosenberg, and J.P. Wilson (1970) The Effective Bandwidth of Cochlear Nerve Fibres. *J. Physiology* 207:62P-63P.

Green, D.M., and B.G. Berg (1991) Spectral weights and the profile bowl. *Quarterly J. of Experimental Psychology. a, Human Experimental Psychology* 43(3)449-58.

Hall, J.W., M.P. Haggard, and M.A. Fernandes (1984) Detection in noise by spectro-temporal pattern analysis. *J. Acoustic Soc. of America* 76(1)50-56.

Hammond, P., and I.M.E. Munden (1990) Areal influences on complex cells in cat striate cortex: Stimulus-specificity of width and length summation. *Experimental Brain Research* 80:135-47.

Heil P, G. Langner and H. Scheich (1992a) Processing of frequency-modulated stimuli in the chick auditory cortex analogue: evidence for topographic representations and possible mechanisms of rate and directional sensitivity. *Journal of Comparative Physiology. a, Sensory, Neural, and Behavioral Physiology* 171(5):583-600.

Heil, P., R. Rajan, and D.R.F. Irvine (1992b) Sensitivity of Neurons in Cat Primary Auditory Cortex to Frequency-Modulated Stimuli. II: Organization of Response Properties along the 'Isofrequency' Dimension. *Hearing Research* 63:135-56.

Hillier, D.A. (1991) *Auditory Processing of Sinusoidal Spectral Envelopes.* Doctoral Thesis. Washington University, St. Louis.

Horst, J.W., E. Javel, and G.R. Farley (1990) Coding of spectral fine structure in the auditory nerve. II. Level-dependent nonlinear responses. *J. Acoust. Soc. Am.* 88(6)2656-81.

Hose B., G. Langner, and H. Scheich (1987) Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? *Brain Research*, 422(2):367-73.

Houtgast, T. (1977) Auditory-Filter Characteristics Derived from Direct-Masking Data and Pulsation-Threshold Data with a Rippled-Noise Masker. *J. Acoust. Soc. Am.* 62(2):409-415.

Houtsma, A.J.M., and J.L. Goldstein (1972) The Central Origin of the Pitch of Complex Tones: Evidence for Musical Interval Recognition. *J. Acoust. Soc. Am.* 51(2):520-529.

Imig, T.J., and H.O. Adrian (1977) Binaural Columns in the Primary Field (A1) of Cat Auditory Cortex. *Brain Research* 138:241-57.

Imig, T.J., W.A. Irons, and F.R. Samson (1990) Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J. Neurophys.* 63(6)1448-56.

Jagadeesh B, H.S.Wheat, and D. Ferster (1993) Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science* 262(5141):1901-4.

Kowalski, N., H. Versnel, and S.A. Shamma (1993) Characteristics of an anterior field in the ferret auditory cortex. *Abstr. Assoc. Res. Otolaryngol.* 16:518.

Langner G, D. Bonke, H. Scheich (1981) Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. *Experimental Brain Research*, 43(1):11-24.

Langner, G., M. Sams, P. Heil, L.K. McEvoy, R. Hari, and A. Ahonen (1994) Periodicity pitch is represented topographically in the human auditory cortex: Evidence by magnetoencephalography. *Gottingen Neurobiology Report 1994 vol II.* 2:388.

Maffei, L., and A. Fiorentini (1973) The visual cortex as a spatial frequency analyser. *Vision Research* 13:1255-67.

Maffei, L., and A. Fiorentini (1976) The unresponsive regions of visual cortical receptive fields. *Vision Research* 16:1131-9.

Mendelson, J.R., and M.S. Cynader (1985) Sensitivity of cat primary auditory cortex (A1) neurons to the direction and rate of frequency modulation. *Brain Research* 327:331-5.

Mendelson, J.R., and K.L. Grasse (1992) A comparison of monaural and binaural responses to frequency modulated (FM) sweeps in cat primary auditory cortex. *Exp. Brain Research* 91:435-54.

Mendelson, J.R., C.E. Schreiner, M.S. Sutter, and K.L. Grasse (1993) Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. *Exp. Brain Research* 94:65-87.

Merzenich, M.M., P.L. Knight, and G.L. Roth (1975) Representation of cochlea within primary auditory cortex in the cat. *J. Neurophys.* 38(1):231-49.

Middlebrooks, J.C., R.W. Dykes, and M.M. Merzenich (1980) Binaural response-specific bands in primary auditory cortex (A1) of the cat: topographical organization orthogonal to isofrequency contours. *Brain Research* 181:31-48.

Miller, J.D. (1989) Auditory-perceptual interpretation of the vowel. *J Acoustic Soc America* 85(5)2114-2134.

Moore, B.C.J. (1982) *An Introduction to the Psychology of Hearing.* Academic Press: New York.

Moore, B.C.J., and M.J. Shailer (1991) Comodulation masking release as a function of level. *J Acoustic Soc America* 90(2)829-35.

Narins, P.M., E.F. Evans, G.F. Pick, and J.P. Wilson (1979) A comb-filtered noise generator for use in audiotry neurophysiological and psychophysical experiments. *Ieee Transactions in Biomedical Engineering* 26(1)43-7.

Newman, J.D., and Z. Wollberg (1973a) Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain Research* 54: 287-304.

Newman, J.D., and Z. Wollberg (1973b) Responses of single neurons in the auditory cortex of squirrel monkeys to variants of a single call type. *Exp. Neurology* 40:821-4.

Pantev, C., M. Hoke, B. Lutkenhoner, and K. Lehnertz (1989) Tonotopic organization of the auditory cortex: pitch vs. frequency representation. *Science* 246(4929): 486-488.

Phillips, D.P. (1988) Effect of tone-pulse rise time on rate-level functions of cat auditory cortex neurons: excitatory and inhibitory processes shaping responses to tone onset. *J Neurophysiology* 59(5)1524-39.

Phillips, D.P., and S.E. Hall (1987) Responses of single neurons in cat auditory cortex to time-varying stimuli: linear amplitude modulations. *Exp. Br. Research* 67:479-92.

Phillips, D.P., S.E. Hall, and J.L. Hollett (1989) Repetition rate and signal level effects on neuronal responses to brief tone pulses in cat auditory cortex. *J. Acoust. Soc. Am.* 85(6)2537-2549.

Phillips, D.P., and D.R.F. Irvine (1981) Responses of single neurons in physiologically defined primary auditory cortex (A1) of the cat: frequency tuning and responses to intensity. *J. Neurophys.* 45(1)48-58.

Phillips, D.P., S.S. Orman, A.D. Musicant, and G.F. Wilson (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. *Hearing Research* 18:73-86.

Raggio, M., and C.E. Schreiner (1994) Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation: I. Intensity dependence of firing rate and response latency. *J. Neurophys.* (in press)

Rajan, R., L.M. Aitkin, D.R. Irvine, and J. McKay (1990a) Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and effects of variations in stimulus parametres. *J. Neurophys.* 64(3)872-87.

Rajan, R., L.M. Aitkin, and D.R. Irvine (1990b) Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J. Neurophys.* 64(3)888-902.

Reale, R.A., and T.J. Imig (1980) Tonotopic Organization in auditory Cortex of the Cat. *J. Comp. Neurology* 192(2)265-91.

Richards, V.M., and L.M. Heller (1991) The detection of a tone added to a narrow band of noise: The energy model revisited. *Quarterly J. of Exp. Psych.* 43:481-501.

Rossing, T.D. (1990) *The Science of Sound.* Addison-Wesley Publishing Company, Reading, Massachusetts.

Schreiner, C.E. (1991) Functional Topographies in the primary audtiory cortex of the cat. *Acta Oto-laryngologica* supp. 491: 7-16.

Schreiner, C.E., B. Calhoun, and D. Keeling (1993) Physiology and topography of cortical neurons explored with vowel-like ripple-spectra. *Abstr. Assoc. Res. Otolaryngol.* 16:161.

Schreiner, C.E., J.R. Mendelson, and M.L. Sutter (1992) Functional topography of cat primary auditory cortex: representation of tone intensity. *Exp. Brain Res.* 92(1): 105-122.

Schreiner, C.E., and J.R. Mendelson (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophys.* 64(5):1442-59.

Schreiner, C.E., and M.L. Sutter (1992) Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J. Neurophysiology* 68(5)1487-1502.

Schreiner, C.E., and J.V.Urbas (1988) Representation of amplitude modulation in the auditory cortex of the cat. II: Comparison between cortical fields. *Hearing Research* 32:49-64.

Shamma, S.A., J.W. Fleshman, P.R. Wiser, and H. Versnel (1993) Organization of response areas in ferret primary auditory cortex. *J. Neurophysiology* 89(2): 367-383.

Shamma S.A. and D. Symmes (1985) Patterns of inhibition in auditory cortical cells in awake squirrel monkeys. *Hearing Research,* 19(1):1-13.

Shamma, S.A., H. Versnel, and N. Kowalski (1994) Responses to rippled complex sound stimuli in primary auditory cortex. *Abstr. Assoc. Res. Otolaryngol.* 17:86.

Shu, Z.G., N.V. Swindale, and M.S. Cynader (1993) *Spectral motion produces an auditory after-effect. Nature* 364(6439)721-3.

Suga, N. (1977) Amplitude spectrum representation in the doppler-shifted-CF pro. ssing area of the auditory cortex of the mustache bat. *Science* 196:64-7.

Suga, N. (1965) Functional properties of auditory neurones in the cortex of echo-locating bats. *J. Physiol.* 181:671-700.

Suga, N., and P.H.S. Jen (1976) Disproportionate tonotopic representation for processing CF-FM sonar signals in the mustache bat auditory cortex. *Science* 194:542-4.

Suga N. and T. Manabe (1982) Neural basis of amplitude-spectrum representation in auditory cortex of the mustached bat. *Journal of Neurophysiology* 47(2):225-55.

Sutter, M.L. and C.E. Schreiner (1990) Two-tone responses of single units in cat primary auditory cortex. *ARO abstracts* 13:221.

Sutter M.L. and C.E. Schreiner (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *Journal of Neurophysiology*, 65(5):1207-26.

Sutter, M.L. and C.E. Schreiner (in press) Topography of intensity parameters in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J. Neurophysiology.*

Wightman, F.L. (1973a) Pitch and stimulus fine structure. *J. Acoust. Soc. Am.* 54(2):397-406.

Wightman, ¯. (1973b) The pattern-transformation model of pitch. *J. Acoust. Soc. Am.* 54(2):407-416.

Winter, P., H.H. Funkenstein (1973) Response properties of auditory cortical cells. *Brain Research* 31(2):368.

Woolsey, C.N., and E.M. Walzl (1942) Topical projection of the nerve fibers from local regions of the cochlea in the cerebral cortex of the cat. *Bull. Johns Hopkins Hosp.* 71:315-44.

Worgotter, F., and U. Th. Eysel (1987) Quantitative determination of orientational and directional components in the response of visual cortical cells to moving stimuli. *Biological Cybernetics* 57:349-55.

Zhang, J. (1990) How to unconfound the directional and orientational information in visual neuron's response. *Biological Cybernetics* 63(2) 135-42.

Table 1
Hypothesis: There is no difference between the RTFs for single units and
multiple units in terms of their best ripple density, bandwidth at 50%, and
modulation index for an α = 0.1.

		count	mean	std. dev.	d.f.	compare variances		compare means		signif.?
						test stat.	crit. value	test stat.	crit. value	
BRD	single	54	1.11	0.86	165	1.47	1.48	0.54	1.65	No
	multiple	113	1.25	1.16						
BW @ 50%	single	49	1.36	0.8	151	1.07	1.52	0.26	1.66	No
	multiple	104	1.29	0.84						
RMI	single	58	0.81	0.16	173	1.49	1.48	0.76	1.65	Yes
	multiple	117	0.71	0.21						

Conclusion: the best ripple density and the bandwidth are not from
significantly different populations. Although the RMI are from different
populations (significantly different variances), the means are not
significantly different.

Table 2
Hypothesis: There is no difference between the RTF filter shapes for
single and multiple units for an α = 0.1.

	band pass	notch	low pass	high pass	mixed
single units	37	6	16	1	16
multiple units	75	8	27	5	41

test statistic = 2.3; critical value = 13.3 (for α > 0.1, d.f. = 8)
Conclusion: the hypothesis is not rejected.

Table 3
Hypothesis: There is no difference between the positive and negative
asymmetries for an α = 0.1.

	count	mean	std.dev.	d.f.	compare variances		compare means	
					test stat.	crit. value	test stat.	crit. value
pos. asymmetry	5	0.352	0.2	10	0.276	1.8	1.29	3.97
neg. asymmetry	7	0.429	0.19					

test statistic = 0.28; critical value = 1.8 (for α > 0.1, d.f. = 10)
Conclusion: the hypothesis is not rejected

Table 4
Hypothesis: There is no difference between the monotonicity and
nonmonotonicity of single and multiple unit populations for an $\alpha = 0.1$.

	monotonic	nonmonotonic	inconclusive
single	3	18	7
multiple	5	7	2

test statistic $= 3.88$; critical value $= 7.77$ (for $\alpha > 0.1$, d.f. $= 4$)
Conclusion: the hypothesis is not rejected

Fig. 1: **Vocalizations:** a) Schematized spectral envelope of a human vowei vocalization. The three formants are marked. b) Actual spectral envelope of a cat vocalization. The two regions of highest energy are marked. Note that the two peaks are nearly one octave apart.

Fig. 2: **Ripple Stimulus:** a) Standard stimulus: harmonic series of 100 to 200 components linearly spaced by the fundamental frequency (F_0), three octaves wide (BW), ripple density (spacing of peaks) of 1 ripple/octave (RD), modulation depth (amplitude of waveform) 30 dB (MD), and a 6dB roll-off to maintain constant energy per octave (T). b) Standard stimulus except ripple density has been changed to 2 ripples/octave. c) Standard stimulus except modulation depth has been changed to 10 dB. Note that the amplitude maxima are the same, but the minima have changed.

Fig. 3: **Placement of Ripple Stimuli** A schematic of a tuning curve shows the ripple stimulus positioned so that the center peak of the ripple stimulus is aligned with the characteristic frequency of the tuning curve. The tuning curve shows both the excitatory region (light gray), and the inhibitory regions (darker gray). With the center peak of the stimulus aligned with the CF of the unit, the phase of the spectral envelope is considered 0°.

Fig. 4: **Ripple Transfer Functions:** a) Post stimulus time histograms showing the responses of units at different ripple densities. The number of responses and the ripple density is shown next to each plot. Although the stimuli were presented in a pseudorandom order, they have been ordered here for ease in interpretation. b) The ripple transfer function plots the number of spikes at each ripple density against the ripple density. The open circles illustrate one set of ten pseudorandom stimuli, the closed circles represent a second set. The line is a two point average of the points. c) A ripple transfer function shows that the best ripple density is 1.5 ripples/octave, the maximum response is 50 spikes, and the minimum response is 7 spikes. The bandwidth is the ripple densities over which the response is greater than half way between the maximum and minimum response: in this case, 2.8 - 1.0, or 1.8 ripples/octave. The modulation index (RMI) is the difference between the maximum and the minimum, divided by the maximum: RMI = (50 - 7)/50, = 0.86.

Fig. 5: **Ripple Transfer Functions:** Twelve different ripple transfer functions showing typical shapes. The plots were determined by using ripple stimuli with modulation depths of 30 dB. The fundamental frequency was constant for each plot. a - f) Typical band-pass filters with the best ripple density ranging from 0.5 to 5.0 ripples/octave. g) Notch filter h) Low pass filter i) Double band pass. These are counted as band pass, and the best ripple density is the overall best response. j) Band pass filter. k) High pass filter. The response at the highest ripple densities never fell below 50% of the maximum response. l) A mixed response or all-pass.

Fig. 6: **Frequency Distributions:** a) Filter Types: Histograms show the number of units that are classified in each particular filter type for single units and multiple units. There was no significant difference between the two populations. b) Best Ripple Density: Histograms showing the number of units that have each of the best ripple densities for single units and multiple units. The bin size is 0.3 ripples/octave. There was no significant difference between the mean and variance for these two populations. c) Bandwidth of

RTFs: Histograms showing the number of units having the specified RTF bandwidths for single units, and multiple units. The bin size is 0.25 ripples/octave. d) Index of Modulation: Histograms showing the number of units having the specified index of modulation for single units, and multiple units. The bin size is 0.05.

Fig. 7: **Phase Shifts:** Responses of six single units as the phase of the spectral envelope is varied. A phase shift of zero corresponds to the center peak of the stimulus being aligned with the CF of unit. The dotted line is a sinusoid fit between the maximum and the minimum response. The modulation depth for all the stimuli was 30 dB. The ripple densities were as follows: a) rd = 1.0 b) rd = 1.0 c) rd = 1. d) rd = 1.8. e) rd = 1.8. f) rd = 1.7.

Fig. 8: **Frequency Distribution for Phase Symmetry Index:** To determine the symmetry of a phase response profile, the symmetry index considers the phase required to reduce the maximum response by 50% as the phase is shifted in the positive and negative directions. The Index is $(\Phi_{pos} - \Phi_{neg})/(\Phi_{pos} + \Phi_{neg})$. The light hatching indicates the phase asymmetry for 11 single units, and the dark hatching, for 1 multiple unit. The bin size is 0.1.

Fig. 9: **Intensity Shifts:** Responses of units as a function of the overall intensity. The stimuli all had a modulation depth of 30 dB, and the following recording site characteristics and ripple densities: a) single unit, rd = 1.0, b) single unit, rd = 1.0, c) single unit, rd = 1.0, d) single unit, rd = 0.9, e) multiple unit, rd = 0.9, f) single unit, rd = 1.0.

Fig. 10: **Frequency Distribution of Dynamic Range of Units:** The positive slope of the intensity response profile indicates the dynamic range of the neurons. To normalize the response, the slope was divided by the difference between the maximum and minimum number of spikes. The dark hatching represents 28 single units while the light hatching represents 15 multiple units. The bin size is 0.01.
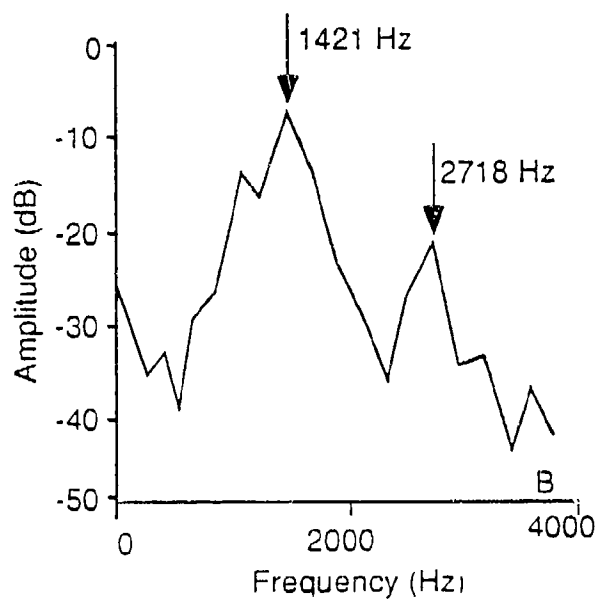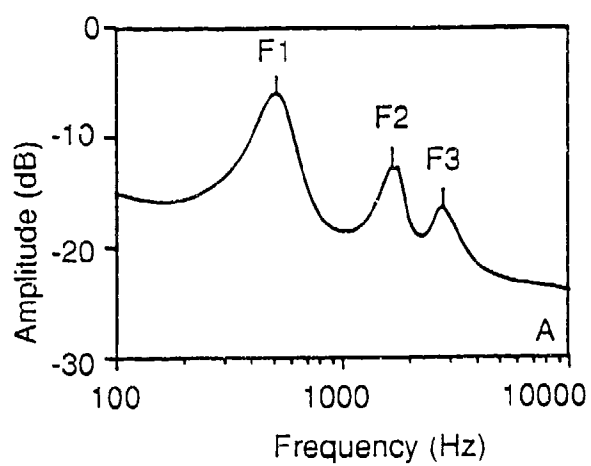
Fig. 11: **Modulation Depth Response Profiles:** Responses of six typical units as the modulation depth of the stimulus is varied. Details of the stimulus and recording site were a) single unit, rd = 1.0, b) multiple unit, rd = 1.0, c) single unit, rd = 1.0, d) single unit, rd = 0.33, e) single unit, rd = 1.5, f) multiple unit. (Closed circles: rd = 1.0 ripples/octave; open circles: rd = 3.0 ripples/octave.)
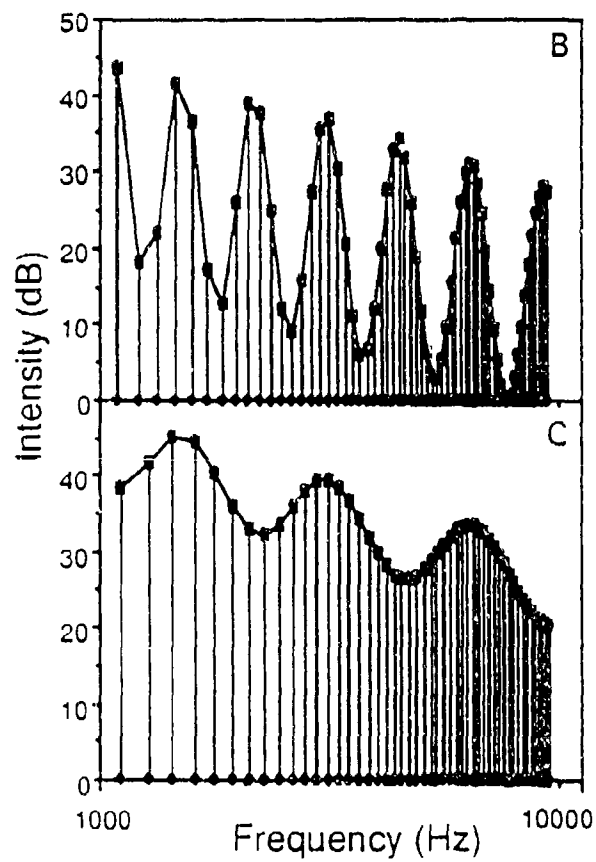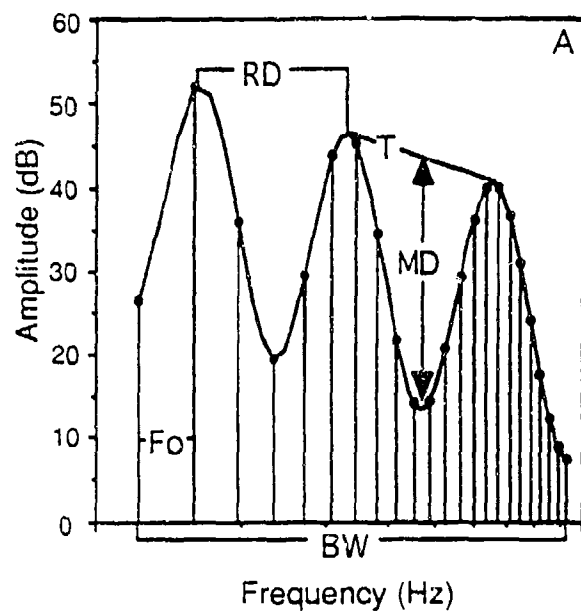
Fig. 12: **Frequency Distributions for Modulation Depth:** The histogram shows the frequency distribution for preferences in the modulation depth. The lightly shaded bars represent single units (n = 6), while the darkly shaded bars represent multiple units (n = 10).

Fig. 13: **Fundamental Frequency Response Profiles:** Change of response magnitude at best ripple density as a function of fundamental frequency. The CFs for these neurons ranged from 4.6 kHz to 7.5 kHz. Details of the recording site and the stimuli: a) multiple unit, cf = 4.6 kHz, rd = 1.0, b) multiple unit, cf = 5.5 kHz, rd = 3.0, c) multiple unit, cf = 5.5 kHz, rd = 1.0, d) multiple unit, cf = 6.9 kHz, rd = 1.0, e) single unit, cf = 6.5 kHz, rd = 1.5, f) single unit, cf = 7.5 kHz, rd = 0.33 ripples/octave.
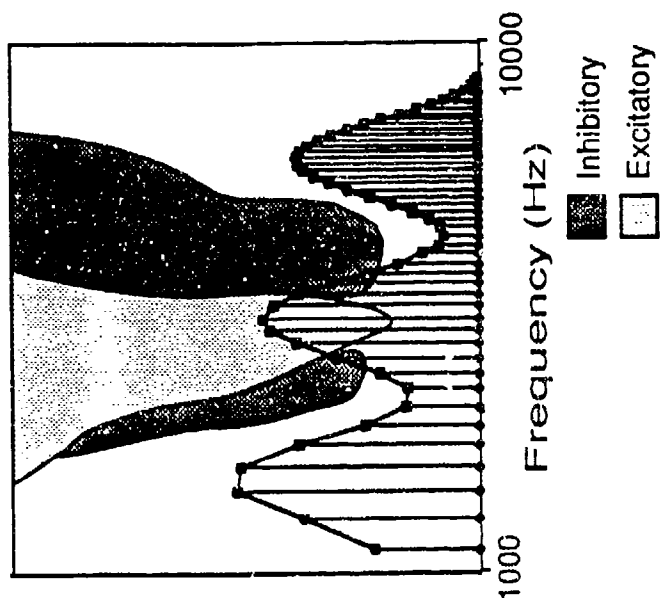
Fig. 14: **Spectral Width Response Profiles:** Six typical spectral width response profiles are shown. a) single unit, rd = 1.66, b) single unit, rd = 1.0, c)

multiple unit, rd = 1.0, d) single unit, rd = 1.0, e) multiple unit, rd = 1.0, f) single unit, rd = 5.0 ripples/octave.

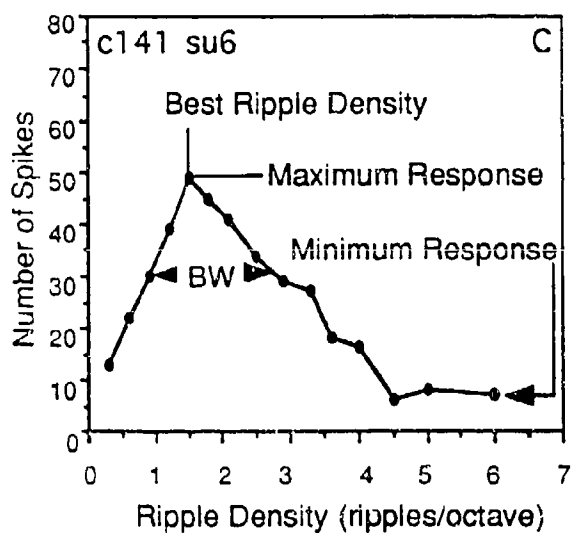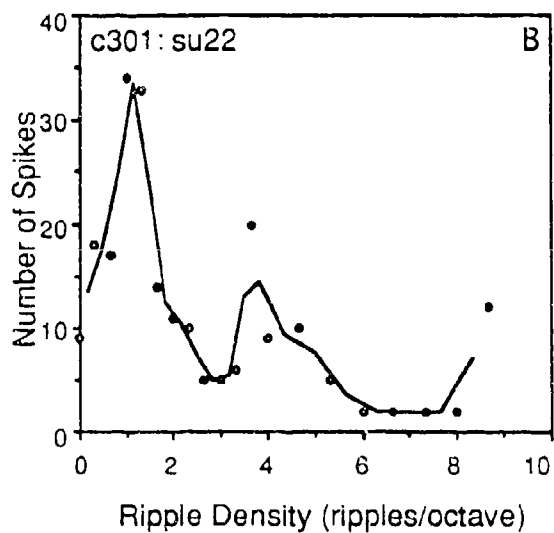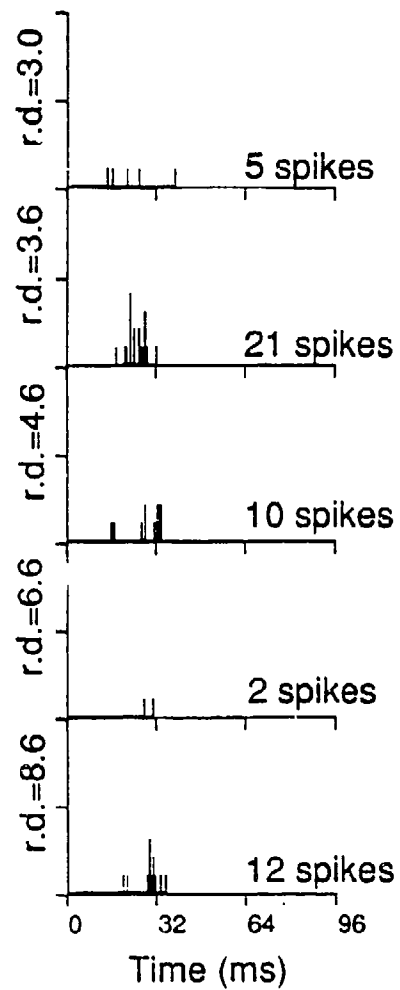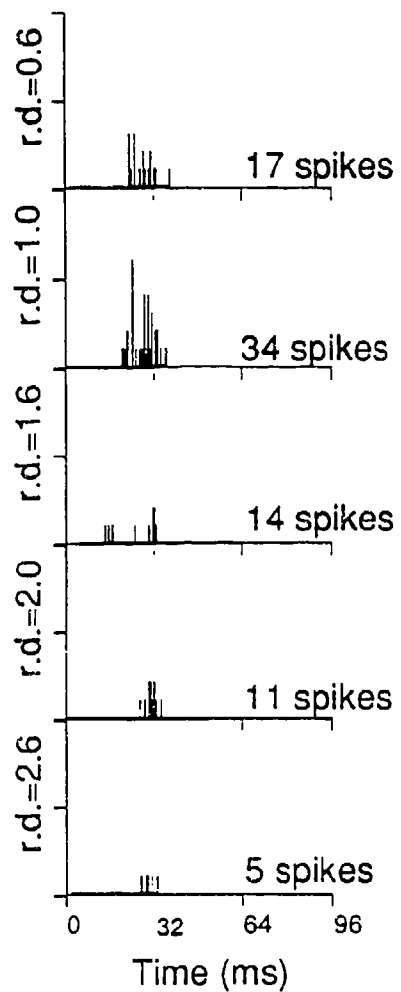schematic from c259 su5d

Frequency (Hz)

1000   10000

■ Inhibitory
□ Excitatory

A

B c301 : su22

C c141 su6
Best Ripple Density
Maximum Response
Minimum Response
BW

Strength of Response (# of spikes)

Ripple Density (ripples/octave)

Strength of Response (# of spikes)

Ripple Phase (degrees)

Symmetry of Phase Response Profiles

Figure showing a histogram with y-axis "Number of Recording Sites" (values 0, 5, 10, 15) and x-axis "Slope of Intensity Response Profile (corrected and normalized)" (values 0.01, 0.05, 0.09, 0.13). Legend: Single Units, Multiple Units.

A

Fundamental Frequency (Hz)

Strength of Response (# of spikes)

Spectral Width (octaves)

c92-1700 su11 A

c92-1700 su18 B

c92-1700 u19 C

c92-1700 su5 D

c92-1700 u7 E

c92-1700 u12 F

TITLE PAGE

Topography of Intensity Tuning in Cat Primary Auditory Cortex: Single-neuron Versus
Multiple-neuron Recordings

M.L. Sutter[1,2] and C.E. Schreiner[1]

[1]Coleman Memorial Laboratory, Department of Otolaryngology, W.M. Keck Center for

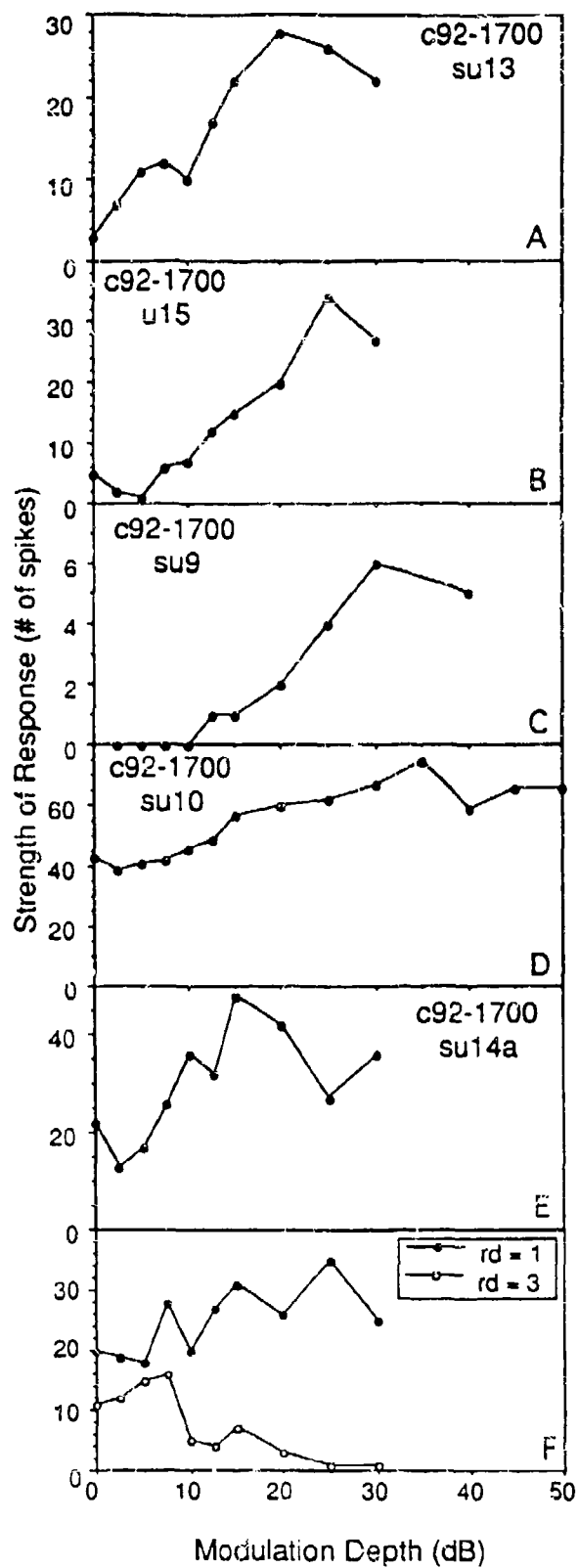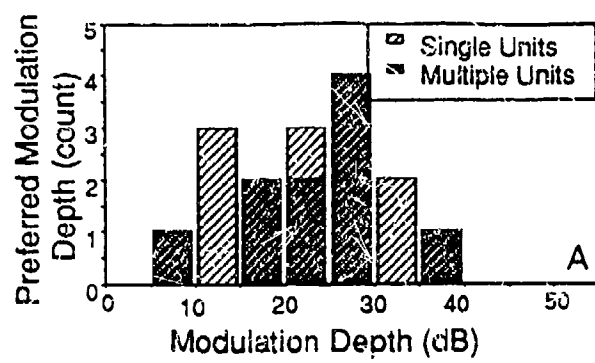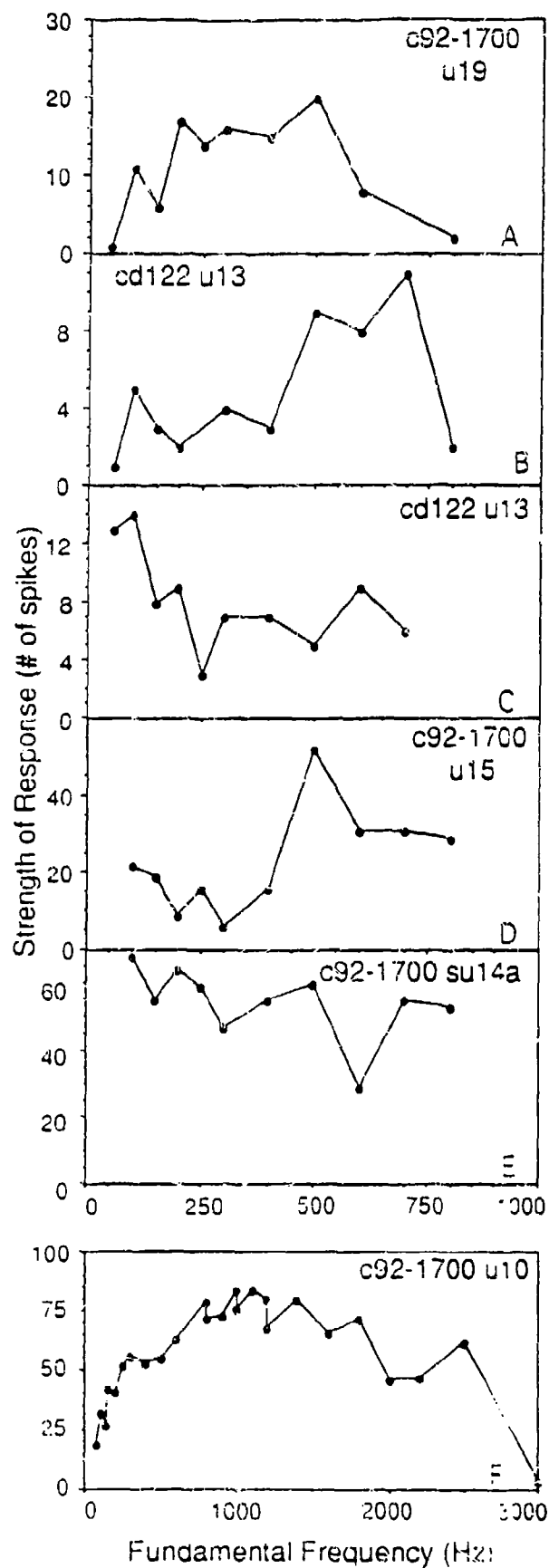Integrative Neuroscience, and Bioengineering Group, University of California San

Francisco, P.O. Box 0732, San Francisco CA 94143-0732

[2]Center for Neuroscience, University of California Davis, Davis CA 95616

Abbreviated Title: Single Unit Intensity Topography in Cat AI

Mailing address for proofs
Dr. Mitchell Sutter
University of California, Davis
Center for Neuroscience
1544 Newton Ct.
Davis, CA 95616
phone: (916) 757- 8849
FAX: (916) 757-8827
e-mail   mlsutter@ucdavis.edu

SUMMARY AND CONCLUSIONS

(1) The spatial distributions of amplitude tuning (monotonicity of rate-level functions) and response threshold of single neurons were studied along the dorsoventral extent of cat primary auditory cortex (AI). To pool data across animals, the multiple unit map of monotonicity was used as a frame of reference. Amplitude selectivity of multiple units is known to vary systematically along isofrequency contours, which run roughly in the dorsoventral direction. Clusters sharply tuned for intensity (i.e., "non-monotonic" clusters) are located near the center of the contour. A second non-monotonic region can be found several millimeters dorsal to the center (Schreiner et al. 1992). The locations of these two non-monotonic regions were used as reference points to normalize data across animals. Additionally, to compare this study to sharpness of frequency tuning results (Schreiner and Mendelson 1990; Schreiner and Sutter 1992; Heil et al. 1992b), multip '
unit bandwidth (BW) maps were also used as references to pool data.

(2) The multiple unit amplitude-related topographies recorded in previous studies were confirmed. Pooled multiple unit maps closely approximated the previously reported individual case maps (Schreiner et al. 1992) when the multiple unit monotonicity or BW_{40dB} map were used as the pooling reference. When the BW_{10dB} map was used as part of the measure, the pooled spatial pattern of multiple unit activity was degraded.

(3) Single neurons exhibited non-monotonic rate-level functions more frequently than multiple units. While common in single neuron recordings (28%), strongly non-monotonic recordings (firing rates reduced by more than 50% at high intensities) were uncommon (8%) in multiple unit recordings. Intermediately non-monotonic neurons (firing rates reduced between 20% and 50% at high intensities) occurred with nearly

equal probability in single neuron (28%) and multiple unit (26%) recordings. The remaining recordings for multiple units (66%) and single units (44%) were monotonic (firing rates within 20% of the maximum at the highest tested intensity).

(4) In ventral AI (AIv), the topography of monotonicity for single units was qualitatively similar to multiple units, albeit single units were on average more intensity selective. In dorsal AI (AId), we consistently found a spatial gradient for sharpness of intensity tuning for multiple units; however, for pooled single units in AId there was no clear topographical gradient.

(5) Response (intensity) thresholds of single neurons were not uniformly distributed across the dorsoventral extent of AI. The most sensitive neurons were consistently located in the non-monotonic regions. The scatter of single neuron intensity threshold was smallest at these locations and increased gradually toward more dorsal and ventral locations.

(6) The existence of a specialized region for near-threshold stimuli along the AIv/AId border is revealed in these experiments. Neuronal recordings in this region are sharply tuned for frequency and amplitude, have low intensity thresholds, have low scatter in characteristic frequency (CF) and threshold, and selectively respond to narrow-band stimuli within 40 dB of the cortical intensity threshold. Non-monotonic neurons have been shown to shift their spike count vs. level functions linearly in response to a continuous noise masker (Phillips et al. 1985; Phillips 1990). Neurons in the ventral non-monotonic region, thus, might serve as fine spectral/amplitude filters which only respond to frequency-banded components with intensities just above the cat's threshold in the presence of background noise.

(7) The results of this study support the parceling of AI into at least two physiologically distinct sub-divisions. The ventral sub-division (AIv) has a complete single unit topographical representation of stimulus intensity. Low intensity signals elicit maximal response at a signal detection region, located at the dorsal extreme of AIv at the AIv/AId border. Neurons respond better to higher intensity signals progressively ventral until the AII border is approached. The dorsal sub-division (AId) is well suited for differential frequency analysis and contains a single unit topography for stimulus bandwidth as previously reported (Schreiner and Sutter 1992).

INTRODUCTION

Understanding the fundamental patterns of representations of stimulus frequency and amplitude are requisite for understanding sound analysis in the auditory system. For mammals, the topographic representation of characteristic frequency (CF) has been well established in primary auditory cortex (AI) (Merzenich et al. 1975; Reale and Imig 1980). Roughly orthogonal to the rostrocaudal tonotopic gradient in the cat, sharpness of frequency tuning is topographically represented (Schreiner and Cynader 1984; Schreiner and Mendelson 1990; Schreiner and Sutter 1992; Heil et al. 1992b). Multiple unit clusters are sharply tuned approximately at the center of dorso-ventral oriented isofrequency contours. Tuning of neuron clusters becomes progressively broader in the dorsal and ventral portions of AI.

Neurons in cat auditory cortex that respond weakly to high intensity stimuli have been described since the earliest microelectrode studies of AI (Erulkar et al. 1956; Evans and Whitfield 1964). The firing rate of such neurons increases above threshold over approximately a 10 to 40 dB range, then decreases with further elevation of stimulus intensity (Brugge et al. 1969; Phillips and Irvine 1981; Phillips et al. 1985; Phillips and Hall 1987). The term "non-monotonic" has been used to characterize these neuronal responses because the general behavior of firing rate as a function of stimulus intensity is non-monotonic (Greenwood and Murayama 1965). Phillips and colleagues (1985) have reported that most non-monotonic neurons are completely unresponsive at high intensities, and thus, their excitatory frequency tuning curves can be described as being circumscribed.

While earlier studies provided preliminary evidence of spatially systematic representation of non-monotonicity in AI (Reale et al. 1979; Phillips et al. 1985), a detailed description of topographies of amplitude characteristics has only recently been published. Multiple unit mapping experiments by Schreiner et al. (1992) have

demonstrated spatial distributions of best-level, threshold, and monotonicity (the sharpness of amplitude tuning) in cat AI. Near and overlapping the region that is sharply tuned for frequency in AI, there is a strongly non-monotonic region. In multiple unit studies, a second non-monotonic region can be located in the dorsal third of AI (e.g., Fig. 1), although the extent and location of the dorsal non-monotonic region varies substantially across animals.

[Fig. 1 about here]

To date, intensity and frequency maps have been obtained using multiple unit techniques (Schreiner and Mendelson 1990; Schreiner et al. 1992; Heil et al. 1992a,b). Deriving single neuron topographies is substantially more difficult. Single neuron experiments yield fewer recorded units because of the difficulties encountering and holding single neurons Resulting single unit maps, thus, either have less dense sampling of the mapped area, a smaller mapped extent, or a smaller set of parameter values than multiple unit maps. Previous attempts to pool single neuron data based on anatomical landmarks yielded results that topographically were only roughly similar to multiple unit studies (Evans and Whitfield 1964, Goldstein et al. 1970). For example, the well-established tonotopic organization of AI was substantially degraded when such a pooling technique was applied (Goldstein et al. 1970). Recent evidence indicates that pooling data based on physiological landmarks is more useful and appropriate (Sutter and Schreiner 1991; Schreiner and Sutter 1992). Comparing single neuron and multiple unit topography is necessary to understand a number of cortical properties. First, we can gain an estimate of the response diversity of the neural elements contributing to local processing. We also can asses how single units contribute to topographies observed for multiple units Finally, we can use the interrelationship of multiple and single unit

2

topographies to try to predict topographical responses to complex stimuli by understanding the local integration of response properties of single units.

In an earlier paper (Schreiner and Sutter 1992), single neuron and multiple unit sharpness-of-frequency tuning maps were compared in AI of the cat at depths between 600 and 1000 microns below the cortical surface. In dorsal AI (AId) multiple unit and single unit topography of BW_{40dB} (Q_{40dB}) were quite similar. However, in ventral AI (AIv), almost all single units were sharply tuned at 40 dB above threshold, even though more ventrally located multiple units showed a tendency towards broader tuning. A nonuniform distribution of local CF scatter across AI was described that contributed to the differences between single and multiple unit responses in AIv.

In this paper, single and multiple unit topographies for the representations of sharpness and sensitivity of amplitude tuning are compared in AI of the cat. Multiple unit maps of sharpness of frequency tuning and monotonicity were used to pool single neuron data across several animals. Similar to the study of excitatory bandwidth, topographical differences in integrative mechanisms along the isofrequency axis of AI are described. Preliminary results from this study have appeared in abstract form (Sutter and Schreiner 1991b).

METHODS

Surgical Preparation

Results presented in this and a previous, related report (Schreiner and Sutter 1992) were obtained in the right hemispheres of 8 young adult cats. Surgical preparation, stimulus delivery and recording procedures for this report are the same as those from a previous study (Schreiner and Sutter 1992).

Briefly, anesthesia was induced with an intramuscular injection of ketamine hydrochloride (10 mg/kg) and acetylpromazine maleate (0.28 mg/kg). After venous cannulation, an initial dose of sodium pentobarbital (30 mg/kg) was administered. Animals were maintained at a surgical level of anesthesia with a continuous infusion of sodium pentobarbital (2 mg/kg/hour) in lactated Ringer's solution (infusion volume: 3.5 ml/hour) and, if necessary, with supplementary intravenous injections of sodium pentobarbital. The cats were also given dexamethasone sodium phosphate (0.14 mg/kg, IM) to prevent brain edema, and atropine sulfate (1 mg, IM) to reduce salivation. The temperature of the animals was monitored with a rectal temperature probe and maintained at 37.5^{o} C by means of a heated water blanket with feedback control.

Three point head fixation was achieved with palatal-orbital restraint (Kopf), leaving the external meati unobstructed. The head holder was attached to the table via magnetic stand. The temporal muscle on the right hemisphere was then retracted and the lateral cortex exposed by a craniotomy (approximately 6-8 mm dorsoventral extent and 3-6 mm rostrocaudal). The dura overlying the middle ectosylvian gyrus was removed, the cortex was covered with silicone oil, and a photograph of the surface vasculature was taken to mark the electrode penetration sites. For maintaining a semi-closed system with a clear view of surface vasculature, occasionally, a wire mesh was placed over the craniotomy and the space between the grid and cortex was filled with a 1% solution of

clear agarose. The agarose-filled grid/chamber helped to diminish pulsation of the cortex and provided a fairly unobstructed view of identifiable locations across the exposed cortical surface.

For cats where histology was performed, at the end of the experiments, the animal was deeply anesthetized and perfused transcardially with saline followed by formalin. Cressyl-violet staining was used to reconstruct electrode position from serial frontal 50 micron sections

Stimulus Generation and Delivery

Experiments were conducted in a double-walled sound-shielded room (IAC). Auditory stimuli were presented via calibrated headphones (STAX 54) enclosed in small chambers that were connected to sound delivery tubes sealed into the acoustic meati [Sokolich 1981, U.S Patent 4251686]. The sound delivery system was calibrated with a sound level meter (Brüel & Kjaer 2209) and a waveform analyzer (General Radio 1521-B). The frequency response of the system was essentially flat up to 14 kHz and did not have major resonances deviating more than +/- 6 dB from the average level. Above 14 kHz, the output rolled off at a rate of 10 dB/octave. Harmonic distortion was at least 55 dB below the primary (depending on the sampling rate and the settings of the antialiasing low-pass filter.)

Tones were generated by a microprocessor (TMS32010; 16 bit D/A converter at 120 kHz; low-pass filter of 96 dB/octave at 15, 35 or 50 kHz). Attenuation was provided by a pair of passive attenuators (Hewlett Packard 350D). The duration of each tone burst was usually 50 ms, except when it was extended to 85 msec for long-latency responses. The rise/fall time was 3 ms. The interstimulus interval was 400 to 1000 ms.

Recording Procedure

Parylene coated tungsten microelectrodes (Microprobe Inc.) with impedances of 1.0-8.5 MOhm at 1 kHz were introduced into the auditory cortex with a hydraulic microdrive (KOPF) remotely controlled by a stepping motor. All penetrations were roughly orthogonal to the brain surface. The recordings reported here were derived at intracortical depths ranging from 600 to 1900 μm, as determined by the microdrive setting. Dimpling was usually less than 100 microns, and thus not a major problem. In several animals, histology indicated that recordings were from layers 3 and 4. Neuronal activity of single units or small groups of neurons (2-6 neurons) were amplified, band pass filtered, and monitored on an oscilloscope and an audio monitor. Multiple unit recordings were employed only to map the sharpness of frequency and amplitude tuning across AI. Recording multiple units allowed for collection of enough data to pre-map 10 to 30 cortical locations in reasonable amount of time (less than eight hours).

Spike activity was isolated from the background noise with a window discriminator (BAK DIS-1). The number of spikes per presentation and the arrival time of the first spike after the onset of the stimulus were recorded and stored (DEC 11/73). The recording window had a duration of 50 to 85 ms, corresponding to the stimulus duration and excluding any offset response.

Data Analysis

[Fig. 2 about here]

From the responses to 675 different frequency/level combinations, an objectively determined frequency response area (Fig 2A) was constructed for every recording site (For more detailed explanation of FRA procedure see Sutter and Schreiner 1991). If the resulting FRA was not well-defined (subjectively determined as no responses for 1/2 of

the points in the tuning curve, or for the level of activity encountered in AI, roughly the standard deviation of responses greater than the mean), the process was repeated with the same 675 stimuli and the resulting evoked activity was added to the first. The process was repeated (up to five times for some multi-peaked units) until a well-defined FRA was obtained. This method has provided statistically reliable characterization of cells based on repeated measures controls (Sutter and Schreiner 1991, Table 3). Response measures, including peak firing rate, were calculated for each stimulus condition of the FRA by weighted averaging with the eight frequency/intensity neighbors as described previously (Sutter and Schreiner 1991).

Spike count versus level functions were derived from each FRA (Fig 2B). The functions were reconstructed by adding action potentials from a 1/4 octave bin centered around the unit's CF (usually 4 different frequencies) over 15 dB (3 levels). This provided at least 12 different stimuli for each tested intensity per repetition of the FRA procedure. Only units which were tested over a 45 dB range or more above minimum threshold were considered for analysis of monotonicity. From the objectively determined single tone FRAs and spike count versus level functions several response properties were measured.

a) Characteristic frequency (CF) = the stimulus frequency with the lowest sound pressure level necessary to evoke neuronal activity.

b) Minimum threshold = lowest intensity associated with stimulus evoked activity in the frequency response area.

c) Bandwidth (BW_{40dB}) = the bandwidth (in octaves) of pure tones to which a cell responds 40 dB above minimum threshold (as measured from the frequency tuning curve) For multi-peaked units the entire bandwidth encompassing the excitatory response (total bandwidth) was measured.

d) Bandwidth (BW_{10}) = the bandwidth (in octaves) of pure tones to which a cell responds 10 dB above minimum threshold (as measured from the frequency tuning

curve). For multi-peaked units the entire bandwidth encompassing the excitatory response (total bandwidth) was measured.

e) best amplitude = The stimulus intensity which elicited the most spikes to any tone.

f) maximum firing rate = The number of spikes, per individual stimulus presentation, at the best amplitude.

g) monotonicity ratio = The number of spikes elicited at the highest intensity level tested divided by the number of spikes elicited at the best amplitude.

[Figure 3 about here]

Single neurons or multiple unit recordings were classified as monotonic if their monotonicity ratio was greater than 0.8 (Figure 3C). Units with ratios between 0.50 and 0.80 were classified as intermediately non-monotonic (Figure 3B). Monotonicity ratios less than or equal to 0.5 were considered strongly non-monotonic (Figure 3A).

Topographical Classification of Single Neurons

[Fig. 4 about here]

Because of the difficulty involved in recording enough parametrically fully characterized single neurons to construct a map for each animal, methods for pooling data across animals were employed. Cytoarchitecture, vasculature and sulcal patterns have historically not been reliable landmarks for pooling (Merzenich et al. 1975), so we used physiological landmarks to directly compare multiple and single unit topography of monotonicity ratio. The pooling method used, therefore, depends on the multiple unit

8

intensity-tuning topography, which in this study was consistent with those of Schreiner et al. (1992).

The axis for the gradients of sharpness of intensity tuning runs roughly in the dorsoventral dimension, orthogonal to the isofrequency axis (Fig. 1). Two regions of sharp intensity tuning (non-monotonic regions) consistently can be found along the dorsoventral axis using the multiple unit technique(Schreiner et al. 1992) The two non-monotonic regions can be visualized as two minima in a plot of monotonicity ratio versus dorsoventral location (Fig. 4).

To pool data across animals, the two minima in the monotonicity ratio versus dorsoventral location plot for each animal were used as normalization points. The monotonicity ratios for all recorded multiple unit clusters for case SUTC16 are shown in Fig. 4A. Initially, the most dorsal recording site was arbitrarily assigned a value of 0 (Fig. 4A, 4B). The two highly non-monotonic regions were approximately 1.0 mm and 3.8 mm from the most dorsal recording site. The data were then re-plotted using a weighted 0.50 millimeter smoothing algorithm (Fig. 4B, see legend). (The convention of placing dashed lines at the two minima of the monotonicity ratio used in Fig. 4B is also used in other figures.) From the resulting smoothed curve, the positions of the minima in monotonicity ratio were extrapolated (e.g., Fig. 4B and Fig. 5 A,C).

Because the ventral and dorsal non-monotonic regions do not have a spatially exact relationship across animals, distance needed to be normalized to use these two locations as pooling landmarks. The distance between the two minima ranged between 2.8 and 3.4 millimeters with a mean of 3.1 (Median 3.05, N=5), and therefore was normalized to "3.0" millimeters (transformation from Fig. 4B to 4C), approximating the median value of 3.05. For the case depicted in Fig. 4, normalization distorted distance by about 7%. This distortion can be seen by directly comparing Fig. 4B and Fig. 4C. In real distance the penetrations from this animal cover slightly less than 5.0 millimeters in dorsoventral extent (Fig. 4B); however, after normalization the penetrations occupy more

than 5 millimeters along the distorted normalized axis (Fig. 4C). After normalization, the dorsal and ventral monotonicity ratio minima were arbitrarily assigned locations of 0.0 and 3.0 millimeters respectively.

[Fig. 5 about here]

The method of pooling data across animals is illustrated in Fig. 5. The un-normalized spatial distributions of single neurons are shown superimposed on the smoothed multiple unit maps (open circles) for 2 cats in panels A and C (A = cat SUTC16, C = cat SUTC12). The multiple unit maps are normalized as previously described, and the locations of single neurons are assigned abscissa values from the normalized coordinate system (Fig. 5B = SUTC16 and Fig. 5D = SUTC12). For case SUTC16 (Fig. 5A and C) single neurons' monotonicity ratios are represented by crosses and for case SUTC12 (Fig. 5B and D) by diamonds. After normalization, single units for both cases have been assigned values relative to the monotonicity map and are directly comparable (compare crosses and diamonds in Fig. 5E to those of 5D and 5B). The same pooling process, as demonstrated in Fig. 5 for cats SUTC12 and SUTC16, was repeated for all other cases.

To enable comparisons of intensity-dependent properties to previous studies, the integrated excitatory bandwidth of multiple unit recordings (Schreiner and Mendelson 1990; Schreiner and Sutter 1992) was also used as a pooling landmark. For pooling relative to bandwidth, 3 measures have been used. The dorsoventral location where the multiple unit responses have the sharpest tuning 40 dB above threshold (BW_{40min}), the dorsoventral location where the multiple unit responses have the sharpest tuning 10 dB above threshold (BW_{10min}), and a composite of the two ($BW_{10/40min}$), namely the average location of the BW_{10min} and BW_{40min}

RESULTS

Results are based on recordings of 108 single neurons recorded from 9 cats. Single neuron recordings were topographically localized with respect to known multiple unit maps. Single neuron population results were compared to 147 multiple units that were recorded in the initial mapping procedure from the same set of experiments.

We found neurons with a wide range of monotonicity ratios (Fig. 3). Throughout AI, single neuron recordings yielded a higher percentage of non-monotonic responses than did multiple unit recordings. Single neurons, particularly those in ventral AI, tended to have a steeper reduction of activity at high intensity levels (i.e., smaller monotonicity ratios) than was observed for multiple unit responses. While the multiple unit topography had two clear non-monotonic regions, the dorsal non-monotonic region was difficult to reconstruct from single unit data. Furthermore, we saw more local topographical scatter in monotonicity ratio and minimum threshold for single neurons than for multiple units.

Differences between single and multiple unit monotonicity ratio topographies can be accounted for by two separate results: (1) a greater scatter in intensity thresholds in monotonic regions. The non-monotonic regions, with lower threshold scatter, contain neurons with similar best amplitudes. (2) Non-monotonic neurons contribute more spikes in topographically non-monotonic regions than in topographically monotonic regions.

Within-Experiment Monotonicity Topography

As reported in an earlier paper (Schreiner et al. 1992), a dorsal and a ventral non-monotonic region were observed in AI. In some cases (4/9), a gradient with increasing sharpness of frequency tuning towards the ventral extreme of AI was found. In these

animals, it appeared that this gradient was leading to a 3rd non-monotonic region in or near AII.

While the intensity dependent response properties of single neurons and multiple units were different in some respects, the locations of the major minima and maxima of the monotonicity ratios of single and multiple unit responses corresponded. Single neurons on average had lower monotonicity ratios, and much wider range of monotonicity ratios at a given dorsoventral location than did multiple units (e.g., Fig. 5). The single neuron topography, was qualitatively similar to the multiple unit topography for the two animals with the largest single unit samples, SUTC12 and SUTC16 (Fig. 5). While similarities between single and multiple units can be seen, the significance is hard to determine because of the large scatter in values for a given dorsoventral location, the small number of samples, and the sharper tuning of single units. Therefore pooling data across animals was necessary for quantitatively verifying the observed single neuron contributions.

Pooled Monotonicity Data

[Fig. 6 about here]

For multiple unit recordings, pooling did not distort the topographical distribution of monotonicity ratio which had a similar shape to that described for individual cases. A weak non-monotonic region near the dorsal non-monotonic reference and a stronger non-monotonic region about 3 mm ventral were observed (Fig. 6A). A third non-monotonic region can be seen at the ventral extreme of AI (from 5-6 mm) for both multiple and single unit topographies. This is caused by the occasionally encountered non-monotonic region near the AI/AII border.

For the pooled data, differences in the details of multiple and single unit response properties can be seen throughout the dorsoventral extent of AI, but in ventral AI the general structure of multiple and single unit topographies were similar. Single units responded with a wider range of monotonicity ratios than multiple units throughout the entire dorsoventral extent of AI (Fig. 6A,B). Single units, also, on average were more sharply tuned for intensity than multiple units. While the wider variability and sharper intensity tuning of single units is apparent, in ventral AI there is a correspondence of the topographical gradients (Fig. 6C). A minimum in monotonicity ratio is located in the ventral non-monotonic region as determined by multiple unit mapping. This minima is accompanied by an ascending gradient in monotonicity ratio for approximately 1.5 mm in either direction for both multiple and single units. In dorsal AI, however there is no correspondence of the pooled topographies. The multiple unit gradient of monotonicity ratio starts descending from a maxima at about 2 millimeters from the dorsal non-monotonic region, then reaches a minima at the center of the dorsal non-monotonic region and reverses (Fig. 6A,C). The topography of single unit monotonicity ratio, however, is relatively flat, except at the dorsal extreme which only has a sample of 4 cells (Fig. 6B). The difference between single and multiple units in dorsal AI indicate that local differences in single unit scatter or the spike contribution of sampled cells are generating the observed multiple unit gradients. The topographical variation of correspondence between single and multiple unit responses demonstrate that there are physiological and/or organizational differences between dorsal and ventral non-monotonic regions.

One might ask whether regions of low monotonicity ratio are a result of a few strongly non-monotonic cells, or of many intermediately non-monotonic cells. From Fig. 6A,B it appears that dorsal AI has fewer strongly non-monotonic neurons than ventral AI, but a similar proportion of intermediately non-monotonic cells. However it is difficult to determine percentages from this figure because there are many cells with a

monotonicity ratio of 1 with overlapping symbols. To address this question
quantitatively, we classified neurons as monotonic (monotonicity ratio > 0.8), strongly
non-monotonic (monotonicity ratio <0.5), or intermediately non-monotonic
(monotonicity ratio between 0.5 and 0.8). A criterion of 0.8 as the monotonic/non-
monotonic cutoff is somewhat arbitrary, and was chosen to be conservative in classifying
a unit as "monotonic".

[Fig. 7 about here]

The percentage of non-monotonic multiple units plotted versus pooled location
shows the spatial distribution that would be expected from the monotonicity ratio
topography (Fig. 7A). Namely, locations with more than 50% non-monotonic neurons
are aligned with the locations that have the lowest monotonicity ratios. There were
however, only a few strongly non-monotonic multiple units. No spatial bin contained
more than 25% of strongly non-monotonic multiple units. The small number of strongly
non-monotonic multiple unit responses roughly inversely followed the spatial distribution
of monotonicity ratio. The percentage of intermediately non-monotonic multiple unit
responses show a more pronounced spatial distribution very similar to the expectations
from the maps shown in Fig. 6C. These results indicate that the spatial distribution of
multiple unit monotonicity ratio is a result of an increase in the percentage of non-
monotonic multiple unit responses at certain locations.

For single neurons, a similar straight-forward explanation of the spatial
distribution of monotonicity did not apply to all cortical locations. While the distribution
of the percentage of non-monotonic single neurons followed the topographical
distribution for monotonicity ratio in ventral AI, this was not as clearly expressed for the
dorsal third of AI (Fig. 6C and Fig. 7B). Although the percentage of non-monotonic
neurons in dorsal AI were similar to the central and ventral region (>60%), there was a

lack of strongly non-monotonic neurons (<10%) in dorsal AI. The high percentage of intermediately non-monotonic neurons in dorsal AI is another property which physiologically distinguishes it from ventral AI. The differences in the sharpness of intensity tuning between neurons of dorsal and ventral non-monotonic regions contribute to the apparent topographical differences between single and multiple unit maps.

Contribution of Firing Rate and Threshold

There are at least two additional contributing factors to how a multiple unit monotonicity topography can be created in dorsal AI, while a similar topography is less evident in the monotonicity of single neurons. One is that non-monotonic neurons contribute more spikes to cluster responses in non-monotonic regions than do monotonic neurons. The other is that, thresholds of neurons might be similar in non-monotonic regions. The resulting frequency tuning curves of non-monotonic single neurons in non-monotonic regions, thus, would be superimposed on each other creating a multiple unit response which is tuned for intensity. By contrast, the threshold scatter in the monotonic regions would obscure the intensity tuning of individual neurons when looking at multiple unit responses.

[Fig. 8 about here]

Part of the difference between multiple and single units in dorsal AI is due to non-monotonic neurons contributing more spikes than monotonic cells. For every neuron, the firing rate at the best amplitude, (FRmax) was calculated. Within each 0.5 or 1.0 mm bin, this firing rate was added for all non-monotonic single neurons (FRmax $_{non-mono}$). FRmax was also added for all monotonic neurons in each bin to arrive at FRmax $_{mono}$. FRmax $_{mono}$ was then divided by FRmax $_{non-mono}$ to arrive at the spike

ratio for any given location (Fig. 8). Ratios greater than 1 indicate that non-monotonic

neurons, on average, contributed more spikes, at best intensity, than monotonic neurons.

The spatial distribution of spike ratios parallels that of multiple unit monotonicity ratios

in the dorsal non-monotonic region. Notice that between 0.0 and 2.0 millimeters the

spatial distribution of single unit spike ratio is different than the that of single unit

monotonicity ratio (Fig. 6). The single unit spike ratio shows a gradient consistent with

the multiple unit topography of monotonicity ratio (Fig. 6 and Fig. 8), while single unit

monotonicity ratio map is flat. Differential contribution of spikes between monotonic and

non-monotonic neurons, therefore, at least partially contribute to the formation of a

multiple unit topographic representation of monotonicity from a relatively non-

topographic single neuron distribution in dorsal AI.

[Figure 9 about here]

        Changes in threshold scatter also contribute to the creation of "monotonic

regions" from underlying non-monotonic cells. Single neuron thresholds were less

scattered in the multiple unit non-monotonic regions than they were in the mapped

monotonic regions. This was apparent in individual cases as well as in the pooled data

(Fig. 9). Notice that in the non-monotonic region near the AI/AII border the scatter was

large, probably reflecting the increased neuron thresholds observed as the AI/AII border

is approached (Schreiner and Cynader 1984). Thresholds in the dorsal and ventral non-

monotonic regions of AI were among the lowest thresholds recorded in each animal (Fig.

9B). The superposition of responses from non-monotonic neurons with low thresholds

contributes to the creation of regions of non-monotonic multiple unit responses.

Relation of Monotonicity Topography to Sharpness of Tuning (BW) Topography

[Figures 10 and 11 about here]

The monotonicity topography is related to the sharpness of frequency tuning topography reported in previous studies (Schreiner and Mendelson 1990, Sutter and Schreiner 1991; Schreiner and Sutter 1992). For pooling data relative to the most sharply tuned "center" of AI (such as in Sutter and Schreiner 1991; Schreiner and Sutter 1992), the measure of the point of alignment chosen is critical. When pooled relative to the location of sharpest frequency tuning as determined by BW_{40dB} (i.e., the bandwidth of pure-tones to which the neuron responds 40 dB above the neuron's minimum threshold), the monotonicity topography is similar to that observed in individual cases. However, when the data were pooled relative to a measure that incorporates the BW_{10dB} distribution the pooled topography was degraded.

Relative to the minima in the BW_{40dB} map, the dorsoventral properties of monotonicity ratio remain similar (Fig. 10A and Figs. 11A, B). The ventral minima is located from 0.0 to 0.5 mm ventral of the sharply tuned BW_{40dB} region and the dorsal non-monotonic region is 2 to 3 mm dorsal to it. As with the monotonicity normalization, the shape of the mean monotonicity ratio vs. location plot is similar in ventral AI, but quite different in dorsal AI (Fig. 10A). However, when the data is pooled by a measure that includes BW_{10dB}, i.e., the bandwidth of pure tones to which the neuron responds 10 dB above threshold, the dorsoventral delineation for amplitude properties degrades (Fig. 10B). This is consistent with the interpretation that non-monotonicity topography is more dependent on high intensity than near threshold properties of sampled neurons.

DISCUSSION

A series of experiments were performed to investigate the topographical distribution of response parameters that contribute to the intensity selectivity of single neurons in cat AI. Consistent with an earlier study (Schreiner et al. 1992), two multiple unit non-monotonic regions spaced about 3 millimeters apart could usually be identified. The ventral non-monotonic region was consistently within a millimeter of BW_{40min}, the location of multiple units most sharply tuned for stimulus frequency 40 dB above threshold. Strongly non-monotonic neurons were concentrated in this ventral non-monotonic region. In dorsal AI, there were many neurons whose firing rates were reduced at high intensities. The magnitude of this reduction on average was smaller than those in the ventral non-monotonic region.

While the ventral non-monotonic gradient could be clearly identified by topographically pooling the mean monotonicity ratios for sampled single units, the dorsal gradient could not be detected in the pooled single unit response data. Differences between single unit and multiple unit representational topographies in dorsal AI could be accounted for by topographical differences in the local threshold scatter of contributing neurons, and by the spike contributions (response strength) of non-monotonic single neurons.

The topographical distribution of multiple unit maps was not lost by the employed pooling methods. While pooled monotonicity maps were sufficiently preserved using BW_{40min} as a pooling landmark, they were substantially degraded when other landmarks, such as BW_{10min} were included.

Methodological Implications

The method of determining the degree of monotonicity in this paper is substantially different from the methods applied by others (e.g., Phillips et al. 1985: Phillips 1985). In earlier work, rate-level functions were collected with 50 repetition peri-time stimulus histograms PSTHs. Maximal firing rates were often below 1 spike per presentation since AI cells strongly habituate to repeated presentation of the same tone. By collecting data for FRA's with a pseudo-random order of 675 different frequency/level combinations, we substantially reduced habituation effects because the average time between similar stimuli was longer. An example of reduced habituation is shown in Fig. 12. The response measured over a 0.1 octave band is comparable to the response to 5 presentations of CF tones (Fig. 12B, inter-stimulus interval = 750 msec). The 5 presentations of CF tones were the first 5 from a stimulation paradigm of 50. When all 50 repetitions were used for analysis, the average response per stimulus presentation was approximately half as strong (Fig. 12), demonstrating that the cell was at least partially habituated for the last 45 stimulus presentations.

[Fig. 12 about here]

Reduced habituation introduced by our method is traded off with problems introduced by adding responses to tones within a 1/4 octave of CF. By adding responses across more than 1 frequency, we may have overestimated the percentage of intermediately non-monotonic neurons. Phillips and colleagues (1985) reported that almost all non-monotonic neurons were completely inhibited at high intensities. The high proportion of intermediate non-monotonic neurons encountered in this study might be partly accounted for by methodological differences. Neurons that do not respond to high

intensity CF tones might still respond to tones whose frequencies are within a quarter octave of CF. The neuron whose FRA is depicted in Fig. 2 and whose spike count versus level functions are displayed in Fig. 12 exemplifies such an "oblique" frequency tuning curve. The neuron s firing rate fails off rapidly at the CF, such that there are no CF responses above 45 dB SPL; however, at intensities greater than 45 dB, there were still responses at frequencies below the CF. The low-frequency response causes the FRA-determined spike count versus level profile to fall off less steeply than the PSTH-determined spike versus level profile (Fig. 12A). Non-monotonic neurons with oblique or tilted frequency tuning curves have been previously shown (e.g., Fig. 3 of Phillips et al. 1985). Many of the intermediate non-monotonic neurons identified in this study had oblique frequency tuning curves. There is no evidence that the intermediate non-monotonic neurons, as defined in this study, had spike versus level functions asymptoting at intermediate response rates. Instead, almost all intermediate non-monotonic neurons showed a downward slope of their spike versus level functions at the highest tested intensities (e.g., Fig 3). Since the intermediately non-monotonic neurons had not reached an asymptote or zero response at the highest intensity tested, we can neither confirm nor rule out the existence of non-monotonic neurons that asymptote at intermediate intensities.

Another methodological consideration is the recording depths encountered in this study. While histology has confirmed that recordings were restricted to layers 3 and 4 in several animals, we cannot rule out the possibility that some recordings extended beyond these boundaries in animals for which no histology was performed. This should not pose a problem because the number of such recordings would have to be minimal and, thus, would be averaged out by pooling. This study did not investigate the effects of depth on single unit topography, but rather attempted to sample from a slab corresponding to layers 3 and 4. While multiple unit maps of intensity are stable with recording depth

(Schreiner et al. 1992), single unit topographies need not be. Therefore, the effect of recording depth is still an open question.

Inter-animal variability

Recent multiple unit mapping studies (Schreiner et al. 1992; Heil et al. 1992b) have reported reproducible topographical distributions of bandwidth and intensity in cat AI. Although, general features of the topography are highly reproducible, there is a high degree of inter-animal variability in the details of the topographies. The details of inter-animal variability are lost in the pooling process. Pooling data across animals, however, allows us to compare the most reliable and reproducible aspects of the topographies across animals for single and multiple units. In these studies, the reliable topographical landmarks are a region of multiple unit responses sharply tuned for frequency and two regions of multiple unit responses sharply tuned for intensity (Schreiner and Mendelson 1990, Schreiner et al. 1992; Heil et al. 1992b). Our ability to obtain pooled results based on these topographical properties is a testament to the reliability of *these aspects* of the multiple unit topography from animal to animal. Our method, however, ignores the inter-animal variability described in these topographies. Therefore, it is important to remember that the goal of these experiments was to compare multiple and single unit responses. In single unit studies, one cannot obtain a large enough extent and fine resolution to measure idiosyncratic inter-animal variability over completely mapped areas (which requires 961 locations to map out a 6 mm by 6 mm grid with 200 micron resolution). Additionally, pooling single and multiple unit data within the same animal will yield irregular maps because of differences in single unit and multiple unit topographies (Schreiner and Sutter 1992). Comparing single unit and multiple unit topographies allow us to gain a better understanding of the basic features of single unit topography underlying the observed multiple unit properties.

An argument can be put forth that our not finding a pooled monotonicity ratio gradient for single units in dorsal AI is due to inter-animal variability. It is possible that within each animal there is a weak single unit gradient for monotonicity ratio. However, there are several arguments against pooling being the sole reason for the absence of a gradient of pooled single unit monotonicity ratio in dorsal AI. First, we do see a gradient for pooled multiple unit maps of intensity parameters in dorsal AI. Therefore, if inter-animal variability is the cause of the lack of a gradient, the variability has to be constrained to single units while not being present in multiple units. Second, we see spatial gradients of spike ratio and threshold for single units in dorsal AI. These results indicate that the lack of spatial gradient or very weak spatial gradient of monotonicity ratio in dorsal AI is not solely a result of pooling data across animals.

Comparison to Previous Studies

The percentage of non-monotonic neurons reported in this paper are consistent with previous results in cat AI. Recent studies from the cat have reported approximately half of AI cells to be non-monotonic. When comparing neuron populations across investigators, the possibility of topographical sampling biases (particularly towards central and ventral AI in recent reports) must be kept in mind (Sutter and Schreiner 1991). Recording from similar cortical depths, Phillips et al.(1985) found that 44% of single neurons (N=61) are strongly non-monotonic, 13% intermediately non-monotonic and 43% monotonic. In a preliminary disclosure, Barone and colleagues (1990) have reported that 40% (data base of 333 neurons) of recorded AI neurons were strongly non-monotonic. In all regions of AI, the reported percentage of non-monotonic neurons in this paper is similar to those of previous studies. In the two millimeters of the dorsoventral extent of AI centered on $BW_{10/40min}$, 33% of neurons are strongly non-monotonic, 28% are intermediately non-monotonic, and 39% monotonic. Dorsal to this

region 10% are strongly non-monotonic and 38% are intermediately non-monotonic. Ventral to this region, 35% are strongly non-monotonic and 19% are intermediately non-monotonic. These results are consistent with previous findings that roughly 40-60% of AI neurons are non-monotonic (see also Figs. 7 and 11). The results from the central and ventral region are consistent with the findings that most non-monotonic neurons in 'classical' AI are strongly non-monotonic (Phillips et al. 1985).

The percentage of non-monotonic neurons is substantially different between AI of the cat and AI of the mustached bat. Most of the neurons in bat AI are intensity selective, although only 25 of 153 (16%) were completely inhibited at high intensities (Suga and Manabe 1982). Of 540 neurons in the same study, only 19 were monotonic. If we assume then that at most 19 of 153 neurons were monotonic, at least 87% of bat AI neurons were non-monotonic, and at least 71% of AI neurons were intermediately non-monotonic. The results of Suga and Manabe indicate that the percentages and strengths of non-monotonic neurons in AI can vary substantially across species.

This interpretation, though, is subject to a severe limitation due to differences between free-field and dichotic stimulus presentation methods. In the bat study, auditory stimuli are presented free-field, while in the cat, stimuli were presented monaurally. The percentage of non-monotonic units found in the medial geniculate body (MGB) and inferior colliculus (IC) indicate that higher percentages of non-monotonic neurons are reported in experiments performed under free-field conditions. In free-field studies of the cat, 80% (Aitken 1991) of IC neurons were non-monotonic. However, recording from MGB under monaural conditions there is a progression in the percentage of non-monotonic neurons from 30% in the anterior portion to 60% in the posterior portion of MGB of the cat (Rodrigues-Dagaeff et al. 1989). Under monaural stimulation conditions in the guinea pig, 24% of IC neurons were non-monotonic (Rees and Palmer 1988). Across these experiments neuronal recordings performed under free-field conditions consistently result in a larger percentage of non-monotonic cells than dichotic

contralateral stimulation. One might expect to see higher percentages of non-monotonic neurons under free-field conditions because of neurons suppressed by the ipsilateral ear with higher thresholds than the excitation threshold in the contralateral ear. Confounding the results further are the different manners in which investigators define the boundaries of brain regions. The relatively few studies of amplitude processing in the central nervous system, coupled with a large variation in methodologies used, make it difficult to conclude whether there are species dependent or brain region dependent differences in the percentages of non-monotonic neurons.

While the percentage of non-monotonic neurons may be different between cat and mustached bat, topographical representations of stimulus intensity can be observed in both species. Besides the cat, functional maps of amplitude related parameters in AI only have been demonstrated in bats. In the mustached bat there is a map of best amplitude that runs roughly orthogonal to the frequency map (Suga 1977; Suga and Manabe 1982). For echolocation, the amplitude of the returned pulse conveys information about object size. The findings of topography of monotonicity and best level in the cat (Schreiner et al. 1992) combined with the single neuron results presented in this study indicate that a topographical representation of intensity parameters might be a general mammalian auditory cortical property. For both species, the amplitude of the stimulus also appears to be encoded orthogonal to the tonotopic axis in AI. While Suga and Manabe did not study the sharpness of amplitude tuning (they studied best-level), there is some indication that there is an orderly organization of degree of monotonicity (see Fig. 10 from Suga and Manabe 1982).

Possible Mechanisms Creating Non-monotonicity

An argument can be put forth that neuronal intensity-tuning is solely a result of the broader frequency response of the basilar membrane at higher intensities (Békésy 1960;

Rhode 1971; Ruggero and Rich 1991) impinging on inhibitory sidebands. The broadening of tuning is due to the mechanical properties of the basilar membrane and effects of stimulus shaping. Therefore, under these circumstances lateral inhibition is converted to intensity tuning. If a splattering of energy into inhibitory sidebands is contributing to the creation of intensity-selective neurons, one would expect to find *inhibitory* sidebands approximately 40 dB below the highest-intensity at which the cell is excited. However, low *excitatory* thresholds, such as those reported for cells in the non-monotonic areas, are unrelated to this potential mechanism. Phillips (1988) has provided evidence that spectral splatter from inhibitory sidebands contributes to intensity tuning by demonstrating that for some non-monotonic cells intensity-tuning becomes sharper with more rapid tone onsets. Based on non-monotonic cells responses to broad-band noise Phillips and Cynader (1985) have also hypothesized that non-monotonicity might arise from inhibitory neurons with similar CFs but higher thresholds. We have recently collected data demonstrating non-sideband high threshold, CF inhibition contributes to intensity tuning (personal observation). These results demonstrate that the auditory system can use sideband (lateral) and/or non-sideband inhibition to create intensity tuning.

While it is probable that sideband inhibition contributes to intensity-tuning , this does not argue that intensity-tuning observed in AI is not functionally relevant. Establishing the significance of intensity-tuning is a behavioral and functional issue, and not only an issue of mechanism. In fact, one would expect the auditory system to exploit spectral splatter for intensity analysis since this is a natural property of sound stimuli.

Functional Significance of Intensity Tuning

Although many studies have been devoted to determining the functional role of the auditory cortical fields, a resolution of distinct functions has not been achieved. There is strong evidence that the AI is involved in sound localization (e.g., Heffner and Masterton

1975; Thompson and Cortez 1983; Jenkins and Merzenich 1984). There is also data which supports the hypothesis that AI is involved in performing vocalization analysis (Gazzaniga et al. 1973; Coslett et al. 1984; Heffner and Heffner 1986a,b, 1989a,b). In bats, there is evidence that auditory cortex performs temporal-spectral analysis (Suga 1989; Simmons et al. 1990).

In addition to sound localization, temporal-spectral, and vocalization analysis, there is a large body of data indicating that auditory cortex is involved in analyzing the intensity of time-varying frequency components comprising auditory scenes (e.g., Simmons et al. 1990; Dear et al. 1993). The presence of intensity maps in AI provides strong circumstantial evidence that AI is contributing to intensity analysis (e.g., Suga and Manabe 1982). While long-duration pure-tone, supra-threshold intensity discrimination seems to be intact in animals with auditory cortical lesions (Swisher 1967, Neff et al. 1975), intensity discrimination appears to be impaired for short-duration tones (Cranford et al. 1982). Additionally data from complex signal discrimination and detection experiments (Heilman et al. 1973; Olsen et al. 1975; Blaettner et al. 1989; Simmons et al 1990), lesion studies (Ferrier 1876, 1889; Muruyama and Kanno 1961; Heffner and Heffner 1986a, 1989a,1990; Jerger et al. 1969, Auerbach et al. 1982), and sound-localization studies (Imig et al. 1991, Rajan et al 1990a;) provide evidence that the auditory cortex is involved in low-intensity signal detection, amplitude-spectrum analysis, signal/noise analysis, and amplitude modulation analysis.

Auditory cortex has been shown to be involved in analyzing many attributes of signals. The physiological organization of auditory cortex is well suited for such multi-dimensional analysis. In AI of the Mustached-bat, cells are tuned to Doppler-shifted-frequency (Suga and Jen 1976), amplitude-spectrum (Suga 1977), echo delays (Fitzpatrick 1993), and to specific intraspecies communication vocalizations (Kanwal et al. 1994; Ohlemiller et al. 1994). Like the bat, AI of the cat has representations of multiple signal parameters (Schreiner et al 1988). The presence of multiple auditory cortical fields and

sub-divisions of these fields argues for multiple functions in auditory cortex. Therefore it is reasonable to suppose that AI can be involved in processing many different signal parameters. Ultimately the only way to address the functional significance of the intensity maps reported in this paper directly is to perform combined behavioral and physiological studies of the role of AI in intensity analysis.

Implications for Sound Localization

Recently there has been a resurgence in interest in sound localization sensitivity of AI neurons (Rajan et al. 1990 a,b; Ahissar et al. 1992). Non-monotonic neurons have been shown to be predominantly directionally selective (Imig et al. 1990). The spatial distribution of minimum thresholds (Fig. 9) indicates that the ventral multiple unit non-monotonic region with its restricted intensity response and threshold range cannot encode intensity-independent location. While the cells in this topographical region would likely be directionally selective, the paucity of monotonic cells that could fire at high intensities indicate that this region can only encode location at low intensities. This, coupled with the fact that monotonic cells are less likely to be directionally selective, make the ventral non-monotonic region a poor candidate for level-tolerant spatial encoding. The monotonic multiple unit regions would be better candidates to serve such a function, since there is a full representation of intensities (including many monaurally non-monotonic single neurons and possibly more free-field non-monotonic neurons). The recent data, however, indicate that a simple intensity-tolerant map of auditory space is unlikely to be found within cat AI.

Possible Purpose of Non-monotonic Regions: Signal Detection?

If the non-monotonic multiple unit regions in AI are not encoding intensity-independent location information, what are they encoding? One possibility is that the strongly non-monotonic neurons are encoding spectral information just above the background noise. The ventral non-monotonic region, which closely overlaps BW_{40min}, could serve this purpose well.

The hypothesis of an area specialized for signal analysis just above background noise is supported by (1) the low thresholds of these neurons (Fig 9), (2) the low scatter in frequency and threshold values, (3) the sharp frequency tuning, (4) the permanent increases in intensity thresholds following AI lesions (Muruyama and Kanno 1961; Heffner and. Heffner 1986a, 1990), and (5) the effect of an increase of background noise on the responses of non-monotonic neurons (Phillips 1985,1990; Phillips and Cynader 1985). Neurons in the ventral non-monotonic region consistently had the lowest thresholds within the isofrequency domain and have the smallest range of threshold values (Fig 9). This implies that the intensity sensitivity is restricted to a small range of low levels. When tones are presented in the presence of continuous (ongoing) background noise, the spike count versus intensity function of non-monotonic neurons shifts as a linear function of the intensity of the noise masker (Phillips 1985,1990; Phillips and Cynader 1985).

Along with the prediction that neurons in the ventral non-monotonic regions are able to shift their spike count versus level functions relative to background noise, these neurons respond to a narrow range of frequencies, and respond weakly, if at all, to broad-band stimuli (Phillips et al. 1985; Schreiner et al. 1988; Schreiner and Mendelson 1990). Accordingly the spectral analysis of signals close to the background signal benefits from a narrow band analysis. Limiting bandwidth enables the neurons to use coherent modulations to detect signals in noise. The ventral non-monotonic region approximately

lines up with BW_{10min}, BW_{40min}, and the lowest threshold region. Topographically, the region near the "physiological center" of AI is highly selective for detecting the frequency and amplitude of narrow-band stimuli just above the background noise, and might be interpreted as a sub-region specialized for signal detection. The narrow-band frequency response of this area and the spatial selectivity of non-monotonic neurons (Imig et al. 1990) combined with their low thresholds indicate that the two non-monotonic regions may play an important role in spectral-spatial processing for stimuli whose intensities are within about 40 dB of the detection threshold of background noise. Note that natural stimuli to which this region would respond must contain at least one narrow-band because most non-monotonic units do not respond well to broad-band stimuli (Phillips et al. 1985, Schreiner and Mendelson 1990).

Parceling of AI

The results of this study raise some questions about the division of AI. Recent papers (Sutter and Schreiner 1991; Schreiner and Sutter 1992) suggested that there are at least two physiologically distinct regions in AI. In dorsal AI (AId) there is a single unit map for stimulus bandwidth. Broadly tuned neurons are located dorsally and there is a gradual progression of narrower tuning toward the dorsoventral center of AI. In ventral AI (AIv), most single units are sharply tuned for frequency. The amplitude tuning described in this paper provides further support for the physiological sub-division of AI into two parts. In ventral AI, where there is no topographical gradient for the bandwidth of single units, there is a single unit map for stimulus intensity. However in dorsal AI, where there is a single unit map for bandwidth, there is no single unit map for stimulus intensity.

While subdividing AI into at least two physiological sub-divisions is supported by the data from this and other studies, the dividing line between AIv and AId remains

uncertain. The implied dividing line from the previous study was the $BW_{10/40min}$ (Q-max in Sutter and Schreiner 1991) as derived from multiple unit mapping of BW_{40dB} and BW_{10dB}. In this study, $BW_{10/40min}$ was not as strong a topographical pooling landmark as was BW_{40dB} or the monotonicity map. The topographical analysis for monotonicity was critically dependent on using high intensity properties of neurons.

With this in mind, the question must be raised whether there is a clear dividing line between dorsal and ventral AI. The answer appears to be that there is no sharp boundary or unambiguous line that can be drawn to determine functionally distinct AI regions. The AIv, AId classification from Schreiner and Sutter (1992) was adequate for studying bandwidth properties. For studying the spatial distribution of a given filter property, the multiple unit topography of that property serves as the best topographical pooling landmark. The BW_{40dB} map probably is the most useful predictor of the location of other response properties studied up to this time. A low-threshold region, with neurons sharply tuned for frequency and amplitude, comprises about plus or minus 1 millimeter from BW_{40min}. Ventral to BW_{40min}, almost all single neurons are sharply tuned for frequency, and there is a single unit map of stimulus intensity. Dorsal to BW_{40min}, there is no clear single unit map of intensity, but there is a gradient for bandwidth. The sharpness of frequency tuning in AId closely follows the multiple unit map, while amplitude tuning does not. The construction of the dorsal non-monotonic region, which is present in multiple unit monotonicity map, but not in pooled single neuron maps, can be accounted for by a combination of threshold scatter and the differential spike contribution of single neurons.

REFERENCES

Ahissar, M., Ahissar, E., Bergman, H. and Vaadia, E. Encoding of sound-source location and movement: activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *J. Neurophysiol.* 67:203-215, 1992.

Aitken, L. Rate level functions of neurons in the inferior colliculus of cats measured with the use of free-field sound stimuli. *J. Neurophyisol.* 65:383, 1991.

Auerbach, S.H., Allard, T., Naeser, M., Alexander, M.P. and Albert, M.L. Pure word deafness. Analysis of a case with bilateral lesions and a defect at the prephonemic level. *Brain* 105:271-300, 1982

Barone, P., Irons, W.A., Clarey, J.C., and Imig, T.J. A comparison of the directional properties of nonmonotonic neurons in the auditory thalamus and primary auditory cortex (AI) of cat. *Soc. Neurosci. Abstr.* 16: 719, 1990.

von Békésy ,G. Experiments in hearing. Translated and edited by E.G. Wever. New York, McGraw Hill, 1960.

Blaettner, U., Scherg, M., and von Cramon, D. Diagnosis of unilateral telencephalic hearing disorders. Evaluation of a simple psychoacoustic pattern discrimination test. *Brain* 112:177-195, 1989.

Brugge, J.F., Dubrovsky, N.A., Aitken, L.M. and Anderson, D.J. Sensitivity of single neurons in auditory cortex of cat to binaural tonal stimulation; effects of varying interaural time and intensity. *J. Neurophysiol.* 32: 1005-1024, 1969.

Coslett, H.B., Brashear, H.R., and Heilman, K.M. Pure word deafness after bilateral primary auditory cortex infarcts. *Neurology* 34:347-352, 1984.

Cranford, J.L., Stream, R.W., Rye, C.V., and Slade, T.L. Detection versus discrimination of brief-duration tones: findings in patients with temporal lobe damage. *Archives of Otolaryngology, Chicago,* 108:350-356, 1982.

Dear, S.P., Simmons, J.A. and Fritz, J. A possible basis for representation of acoustic
    scenes in auditory cortex of the big brown bat. *Nature* 364:620-623.

Erulkar, S.D., Rose, J.E. and Davies, P.W. Single unit activity in the auditory cortex of
    the cat. *Johns Hopkins Hospital Bulletin* 39: 55-86, 1956.

Evans, E.F. and Whitfield, I.C. Classification of unit responses in the auditory cortex of
    the unanesthetized and unrestrained cat. *J. Physiol. (London)* 171: 476-493, 1964.

Ferrier, D. Functions of the Brain. Dawsons of Pall Mall: London, 1876.

Ferrier, D. Schafer on the temporal and occipital lobes. *Brain* 11:7-30, 1889.

Fitzpatrick, D.C., Kanwal, J.S., Butman, J.A. and Suga, N. Combination-sensitive
    neurons in the primary auditory cortex of the mustached bat. J. Neuroscience 13:
    931-940, 1993

Gazzaniga, M.S., Glass, A.V., and Sarno, M.T. Pure tone word deafness and hemispheric
    dynamics: a case history. Cortex 9:136-143, 1973

Goldstein, M.H., Abeles, M. and Daly, R.L. Functional architecture in cat primary
    auditory cortex: tonotopic organization. *J. Neurophysiol.* 33: 188-197, 1970.

Greenwood, D.D. and Maruyama, N. Excitatory and inhibitory response areas of
    auditory neurons in the cochlear nucleus. *J. Neurophysiol.* 28: 863-892, 1965.

Heffner, H.E. and Heffner, R.S. Effect of bilateral auditory cortex lesions on absolute
    threshold in Japanese macaques. *J. Neurophysiol.* 64:191-205., 1990.

Heffner, H.E. and Heffner, R.S. Effect of restricted cortical lesions on absolute
    thresholds and aphasia-like deficits in Japanese macaques. *Behav. Neurosci.*
    103:158-169, 1989a

Heffner, H.E. and Heffner, R.S. Cortical deafness cannot account for the inability of
    Japanese macaques to discriminate species-specific vocalizations. *Brain Lang.*
    36:275-285, 1989b

Heffner, H.E. and Heffner, R.S. Hearing loss in Japanese macaques following bilateral auditory cortex lesions. *J. Neurophysiol.* 55:256-271, 1986a.

Heffner, H.E. and Heffner, R.S. Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques *J. Neurophysiol.* 56:683-,701 1986b.

Heffner, H. and Masterton, B. Contribution of auditory cortex to sound localization in the monkey (*Macaca mulatta*). *J. Neurophysiol.* 38: 1340-1358.

Heil, P., Rajan, R. and Irvine, D.R.F. Sensitivity of neurons in cat primary auditory cortex to tones and frequency modulated stimuli: I: Effects of variation of stimulus parameters. *Hearing Research* 63:106-134, 1992a.

Heil, P., Rajan, R. and Irvine, D.R.F. Sensitivity of neurons in cat primary auditory cortex to tones and frequency modulated stimuli: Organization of response properties along the 'isofrequency' dimension. *Hearing Research* 63:134-156, 1992b

Heilman, K.M., Hammer, L.C. and Wilder, B.J. An audiometric defect in temporal lobe dysfunction. *Neurology* 23:384-386, 1973

Imig, T.J., Irons, W.A. and Samson, F.R. Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high frequency primary auditory cortex. *J. Neurophysiol.* 63: 1448-1466, 1990.

Jenkins, W.M. and Merzenich, M.M. Role of cat primary auditory cortex for sound localization behavior. *J. Neurophysiol.* 52: 819-847, 1984.

Jerger, J., Weikers, N.J., Sharbrough, F.W. and Jerger, S. Bilateral lesions of the temporal lobe. A case study. *Acta Otolaryngol. Suppl.* 258:1-51, 1969

Kanwal, J.S., Ohlemiller, K.K., and Suga, N. Communication sounds of the mustached bat: classification and multidimensional analysis of call structure. Abstracts of the 16th midwinter meeting of the Association for Research in Otolaryngology: Abstract # 442, 1994

Merzenich, M.M., Knight, G.L. and Roth, G.L. Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38:231-249, 1975.

Muruyama, N. and Kanno, Y. Experimental study of functional compensation after bilateral removal of auditory cortex in cats. *J. Neurophysiol.* 24:193-202, 1961.

Neff, W.D., Diamond, I.T., and Casseday, J.H. Behavioral studies of auditory discrimination: central nervous system. In: *Handbook of Sensory Physiology. Auditory System.* Ed. by W.D. Keidel and W.D. Neff, New York: Springer-Verlag, vol. V, p. 2, pp. 307-400, 1975.

Ohlemiller, K.K., Kanwal, J.S., and Suga, N. Do cortical auditory neurons of the mustached bat have a dual function for processing biosonar signals and communication sounds? Abstracts of the sixteenth winter meeting of Association for Research in Otolaryngology: Abstract # 441, 1994

Olsen, W.O., Noffsinger, D. and Kurdziel, S. Speech discrimination in quiet and white noise by patients with peripheral and central lesions, *Acta Otolaryngol.* 80:375-382, 1975.

Phillips, D.P. Neural representation of sound amplitude in the auditory cortex: effects of noise masking. *Behavioral Brain Research* 37:197-214, 1990.

Phillips, D.P. Effect of tone-pulse rise time on rate-level functions of cat auditory cortex neurons: excitatory and inhibitory processes shaping responses to tone onset. *J. Neurophysiol.* 58: 1524-1539, 1988

Phillips, D.P. Temporal response features of cat auditory cortex neurons contributing to sensitivity to tones delivered in the presence of continuous noise. *Hearing Research* 19:253-268, 1985.

Phillips, D.P. and Cynader, M.S. Some neural mechanism in the cat's auditory cortex underlying sensitivity to combined tone and wide-spectrum noise stimuli. *Hearing Research* 18: 87-102, 1985.

Phillips, D.P. and Hall, S.E. Responses of single neurons in cat auditory cortex to time-varying stimuli: linear amplitude modulations. *Exp. Brain Res.* 67: 479-492, 1987.

Phillips, D. and Irvine, D.R.F. Responses of single units in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and response to intensity. *J. Neurophysiol.* 45(1):48-58, 1981.

Phillips, D.P., Orman, S.S., Musicant, A.D. and Wilson, G.F. Neurons in the cat's primary auditory cortex distinguished by their response to tones and wide-spectrum noise. *Hearing Research* 18:73-86, 1985.

Rajan, R., Aitken, L.M., Irvine, D.R.F and McKay, J. Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. *J. Neurophysiol.* 64: 872-901, 1990a.

Rajan, R., Aitken, L.M. and Irvine, D.R.F. Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *J. Neurophysiol.* 64: 888-902, 1990b.

Reale, R.A. and Imig, T.J. Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.* 192:265-291, 1980.

Reale, R.A. Imig, T.J. and Sinex, D.G. Rate intensity functions of single neurons located with binaural suppression columns of cat primary auditory cortex. *Soc. Neurosci. Abstr.* 5:29, 1979.

Rees, A. and Palmer, A.R. Rate-intensity functions and their modifications by broadband noise for neurons in the guinea pig inferior colliculus. *J. Acoust. Soc. Am.* 83:1488-1498, 1988.

Rhode, W.S. Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *J. Acoust. Soc.* 49:1218-1231, 1971.

Rodrigues-Daegaeff, C., Simm, G., De Ribaupierre, Y., Villa, F., De Ribaupierre, F. and Roullier, E.M. Functional organization of the ventral division of the medial geniculate body of the cat: evidence for a rostro-caudal gradient of response properties and cortical projections. *Hearing Research* 39:103-126, 1989.

Ruggero, M.A., and Rich, N.C. Application of a commercially-manufactured Doppler-shift laser velocimeter to the measurement of basilar-membrane vibration. *Hearing Research* 51:215-230, 1991.

Schreiner, C.E. and Cynader, M.S. Basic functional organization of secondary auditory cortical field (AII) of the cat. *J. Neurophysiol.* 51(6):1284-1305, 1984.

Schreiner, C.E., Mendelson, J.R., Grasse, K. and Sutter, M.L. Spatial distribution of basic response properties in cat primary auditory cortex. *Assoc. Otolaryngol. Abstr.* 11:36, 1988.

Schreiner, C.E. and Mendelson, J.R. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophysiol.* 64: 1442-1460, 1990.

Schreiner, C.E., Mendelson, J.R. and Sutter, M.L. Functional topography of cat primary auditory cortex: representation of tone intensity *Exp. Brain Res.* 92:105-122, 1992.

Schreiner, C.E. and Sutter, M.L. Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple neuron recordings. *J. Neurophysiol.* 68:1487-1502, 1992.

Suga, N. Auditory combination sensitivity and speech processing: complex-sound processing by combination sensitive neurons. In Auditory Function : Neurobiological Bases of Hearing. Ed. by G.W. Edelman, W.E. Gall, W.M. Cowan. N.Y.: John Wiley and Sons, pp. 670-720, 1989.

Suga, N. Amplitude-spectrum representation in the Doppler-shifted-CF processing area of the auditory cortex of the mustache bat. *Science 196: 64-67, 1977.*

Suga, N. and Jen, P.H.-S. Disproportionate tonotopic representation for processing species-specific CF-FM sonar signals in the mustache bat auditory cortex. *Science* 194:542-544, 1976.

Suga, N. and Manabe, T. Neural basis of amplitude-spectrum representation in auditory cortex of the mustached bat. *J. Neurophysiol.* 47: 225-255, 1982.

Sutter, M.L and Schreiner, C.E. Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65:1207-1226, 1991.

Sutter, M.L. and Schreiner, C.E. Topography of intensity encoding by single neurons in cat primary auditory cortex. *Soc. Neuroscience Abstr.* 17:1484, 1991b.

Swisher, L. Auditory intensity discrimination in patients with temporal lobe damage. *Cortex* 3:179-193, 1967.

Thompson, G.C. and Cortez, A.M. (1983). The inability of squirrel monkeys to localize sound after unilateral ablation of auditory cortex. Behav. Brain Res. 8:211-216.

Figure 1:

Cartoon of multiple unit sharpness of amplitude (A) and frequency (B) tuning maps from previous studies (Schreiner and Mendelson 1990, Schreiner et al. 1992). Monotonicity map (A) has two non-monotonic (dark) regions sandwiched between monotonic (light) regions. The sharpness of frequency tuning (Q40dB) map has a sharply tuned region (dark) which gradually gives way to broader frequency tuning (light). The ventral non-monotonic region roughly lines up with the sharply tuned region (dashed line). Both maps gradients are oriented roughly parallel to the isofrequency domain. SSS: suprasylvian sulcus; PES: posterior ectosylvian sulcus; AES: anterior ectosylvian sulcus.

Figure 2:

One-presentation FRA (A) and corresponding spike count vs. level function (B) for a non-monotonic neuron. Gap in dotted background in (A) represents 1/4 octave band over which spikes were counted. Two spikes on the low frequency side at 55 dB were outside of the 1/4 octave band and thus were not counted for the spike count vs. level plot. The maximum of 0.8 spikes per repetition seems low because the cell was more narrowly tuned than the 1/4 analysis bin used (see Fig. 12 for more discussion of this methodological effect). The monotonicity ratio is calculated by dividing the response at the highest tested intensity, 72 dB, by the largest response. The spike count vs. level plot incorporates data from another 1-repetition FRA (not shown) from the same cell which covered from 2 to 72 dB SPL. Also notice that the frequency range in this example is high resolution (90 frequencies).

Figure 3:

Spike count vs. level functions for six neurons. All spikes from an FRA are counted over 15 dB (3 levels, shown by horizontal lines connecting the 3 data collects for each point) and 1/4 octave (5 frequencies). Some are based on more than 1 repetition of

the FRA procedure, and/or were excited by less than a 1/4 octave frequency range. The monotonicity ratio is calculated by dividing the response at the highest tested intensity by the largest response.

Figure 4:

Spatial distribution of monotonicity ratio within isofrequency domain (7-9 kHz) for cat SUTC12. In (A) the multiple unit data points are shown. The result of a weighting average spatial smoothing algorithm on the data points is shown in (B). Clusters with the same dorsoventral coordinate were assigned weights of 1.0. Recording sites within 0.10 millimeters of each other were assigned weights of 0.75. Multiple units between 0.11 and 0.25 millimeters were assigned weights of 0.5, and neighbors between 0.26 and 0.50 millimeters were assigned weights of 0.25. The weighted average was calculated to give a "smoothed" value for each data point. The two minima in the function are identified by dotted lines. The distance between the minima then was normalized to "3 millimeters" (C) which in this case expanded the real distances by 7%. Notice that in (B) the map is less than 5 millimeters but in (C) it is more than 5. For all graphs the ventral non-monotonic region is assigned the value "3" and the dorsal non-monotonic region is assigned the value "0".

Figure 5:

Illustrative example of pooling method using the two cases shown in Fig 4. The multiple unit maps (connected filled circles) with each single (crosses for cat SUTC16 and open diamonds for case SUTC12) and multiple unit (open circles) are shown for two cases (A and C). Single unit points (same symbols as in A and C) are displayed after normalization of locations to "3 normalized millimeters" from the multiple unit map (B and D). Single units from the 2 cats are superimposed with their common normalized

coordinate system (E). Notice that the point in AII (C and D) is not included in final
pooled data

Figure 6:

Scattergram of pooled monotonicity ratio data for (A) multiple units, and (B) single units. For (A) the dashed line connecting filled rectangles displays the mean monotonicity ratios of multiple units averaged over 0.5 or 1.0 millimeter bins; rectangles are located at the center of the bins. One half millimeter bins were used if at least ten neurons were located within the bin, otherwise 1.0 millimeter bins were used. The numbers of recordings in each bin are displayed at the top of each graph. A similar curve for single neurons is shown in Fig. 6B. Mean binned results for single and multiple units with magnified monotonicity ratio scale (C). For all plots vertical dashes lines represent the normalization points used, i.e., center of the dorsal and ventral non-monotonic regions as determined by multiple unit mapping.

Figure 7:

Percentage of strongly (black filled) and intermediately (hatched) non-monotonic neurons for topographically pooled multiple (A) and single (B) units. Number of units pooled for each bin is shown by n value on top of each plot.

Figure 8:

Ratio of spikes contributed from non-monotonic units and monotonic units. Maximum spike count per stimulus to tones (FRmax) were calculated for each neuron. Within each bin, FRmax was added for all non-monotonic neurons, and this sum was then divided by the equivalent measure for monotonic neurons. Notice that a maximum of the ratio occurs both in the dorsal and ventral non-monotonic regions.

Figure 9:

 Threshold data pooled across animals. Zero and three millimeters correspond to two non-monotonic regions. The line drawn at 3 millimeters for all other figures has been omitted to allow for better viewing of the data. Threshold was referenced such that for each animal 0 dB corresponded to the lowest multiple unit threshold encountered. Notice the minima in threshold scatter at the two non-monotonic regions and the threshold minimum at the ventral non-monotonic region. Unclassifiable Monotonicity (open boxes in A) refers to cells in which monotonicity ratio could not be calculated because less than 50 dB of recording range was achieved.

Figure 10:

 Mean binned monotonicity ratio pooled relative to BW40min and BW10/40 min as described in Schreiner and Sutter 1992.

Figure 11:

 Percentage of strongly (black) and intermediately (stippled) non-monotonic single (A,C) and multiple (B,D) units. Topographical pooling performed relative to BW40min and BW10/40min.

Figure 12:

 Habituation effects of the PSTH method for neuron whose FRA is depicted in Fig. 2. Comparison of spike count vs. level plot derived from 50 repetition CF tone PSTHs with measures derived from FRA's (A). Notice that the 0.25 octave window used underestimates spikes per presentation. This is because the neurons FTC is less than 1/4 octave wide (BW is approximately 0.15 octaves). The 50 repetition PSTH also underestimates the number of spikes per presentation as compared to the random order FRA method. This is probably due to habituation. When only the first 5 repetitions of

each PSTH are counted (B), the habituation effect disappears. The 0.1 octave FRA method and the PSTH method now yield comparable results. Notice, though, that the descending branch (intensities greater than 20 dB) is less sharply sloped for the FRA method as compared to PSTHs. This is because at high intensities, to which the CF response is substantially reduced, off-CF responses are still present (see Fig. 2).
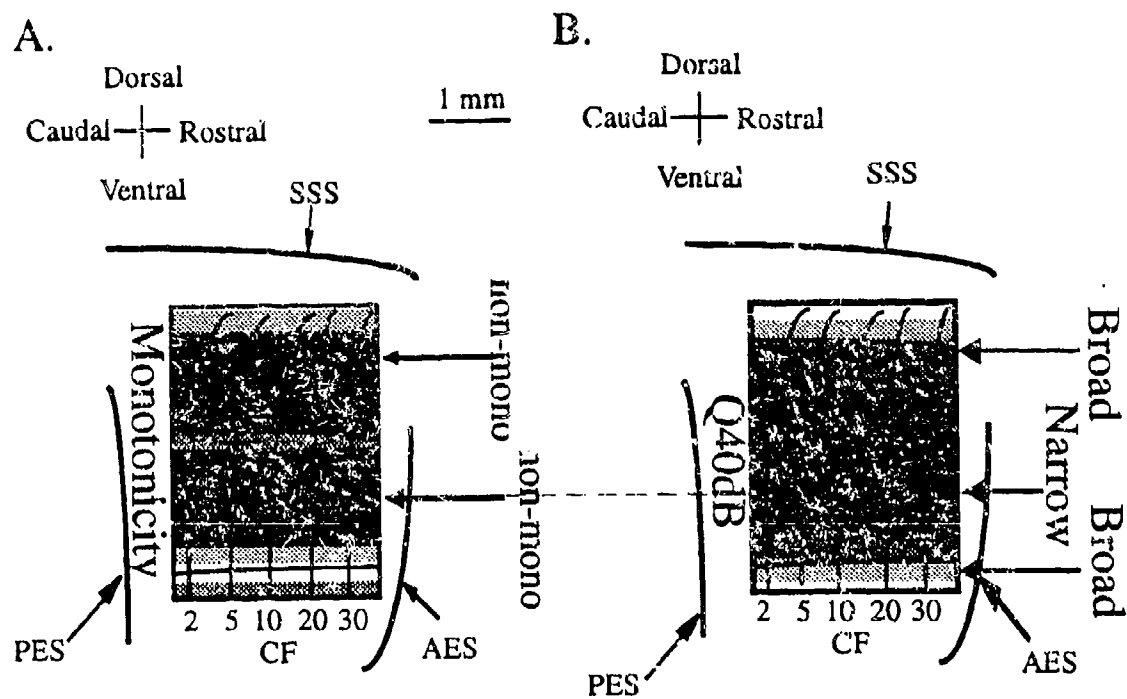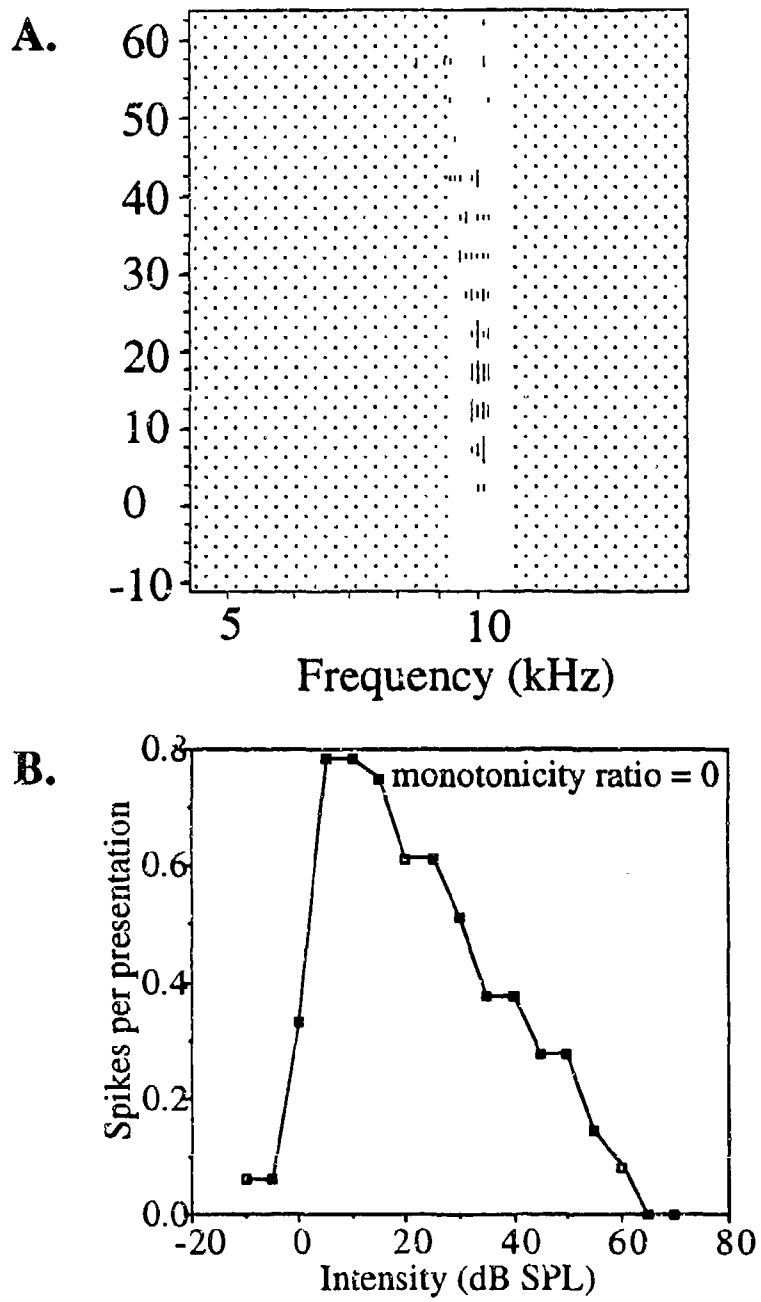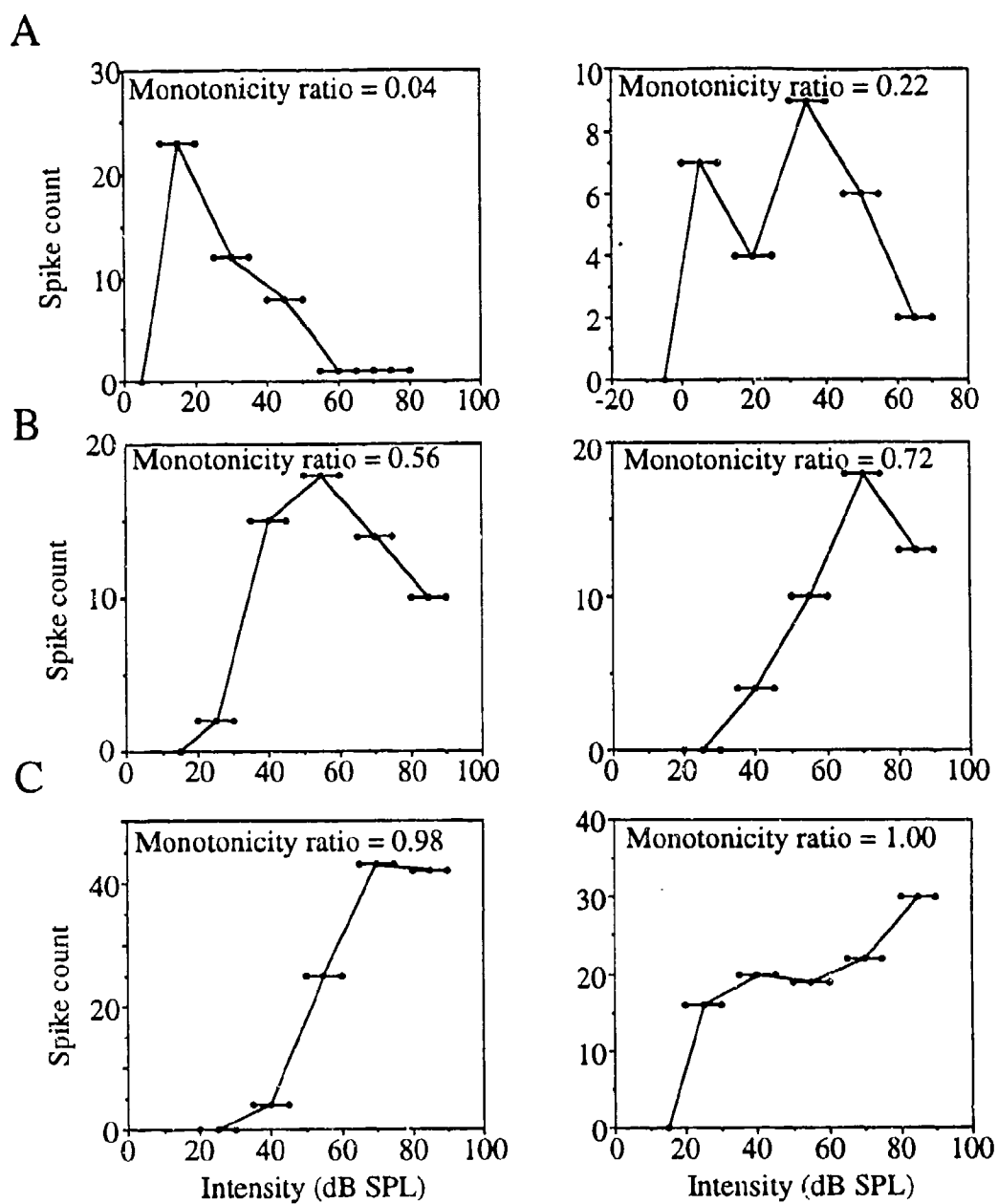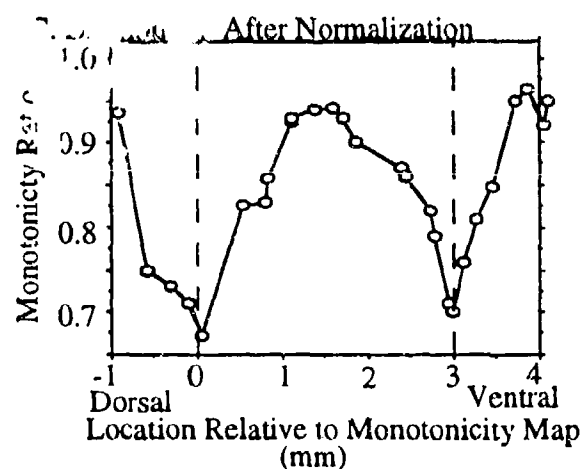
Figure 1

A.

B. monotonicity ratio = 0

Figure 2

Figure 3

A. Multiple Units

B. Multiple Units with Spatial Smoothing

After Normalization

Figure 4

A. SUTC16 unnormalized

B. SUTC16 normalized

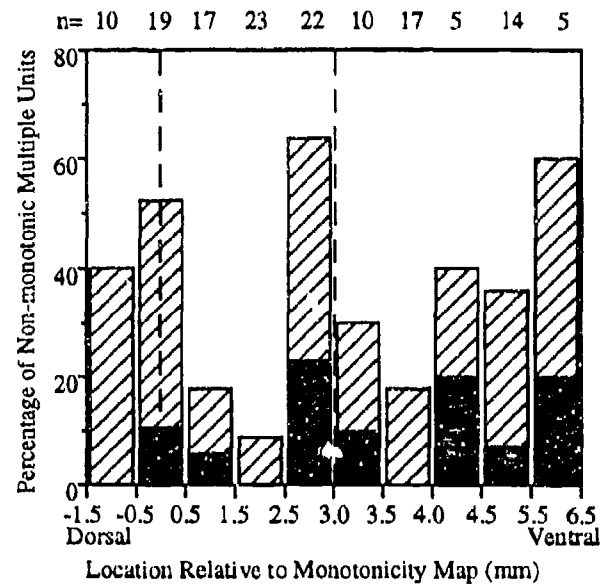C. SUTC12 unnormalized

D. SUTC12 normalized

E. Pooled From SUTC16 and SUTC 12
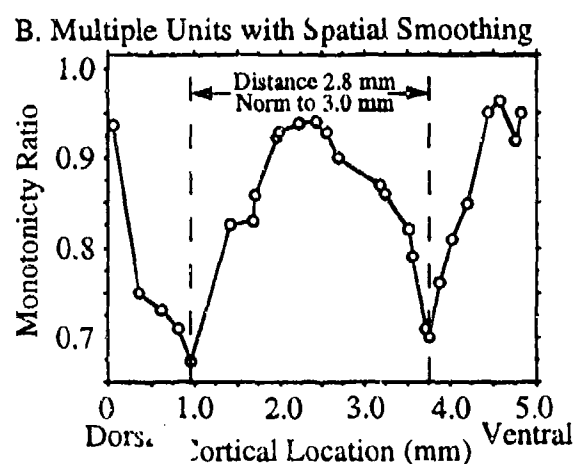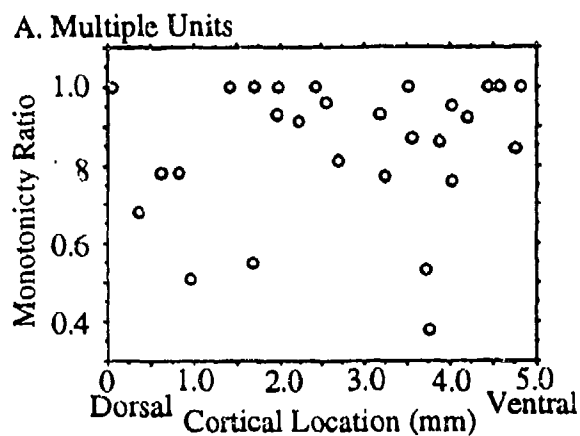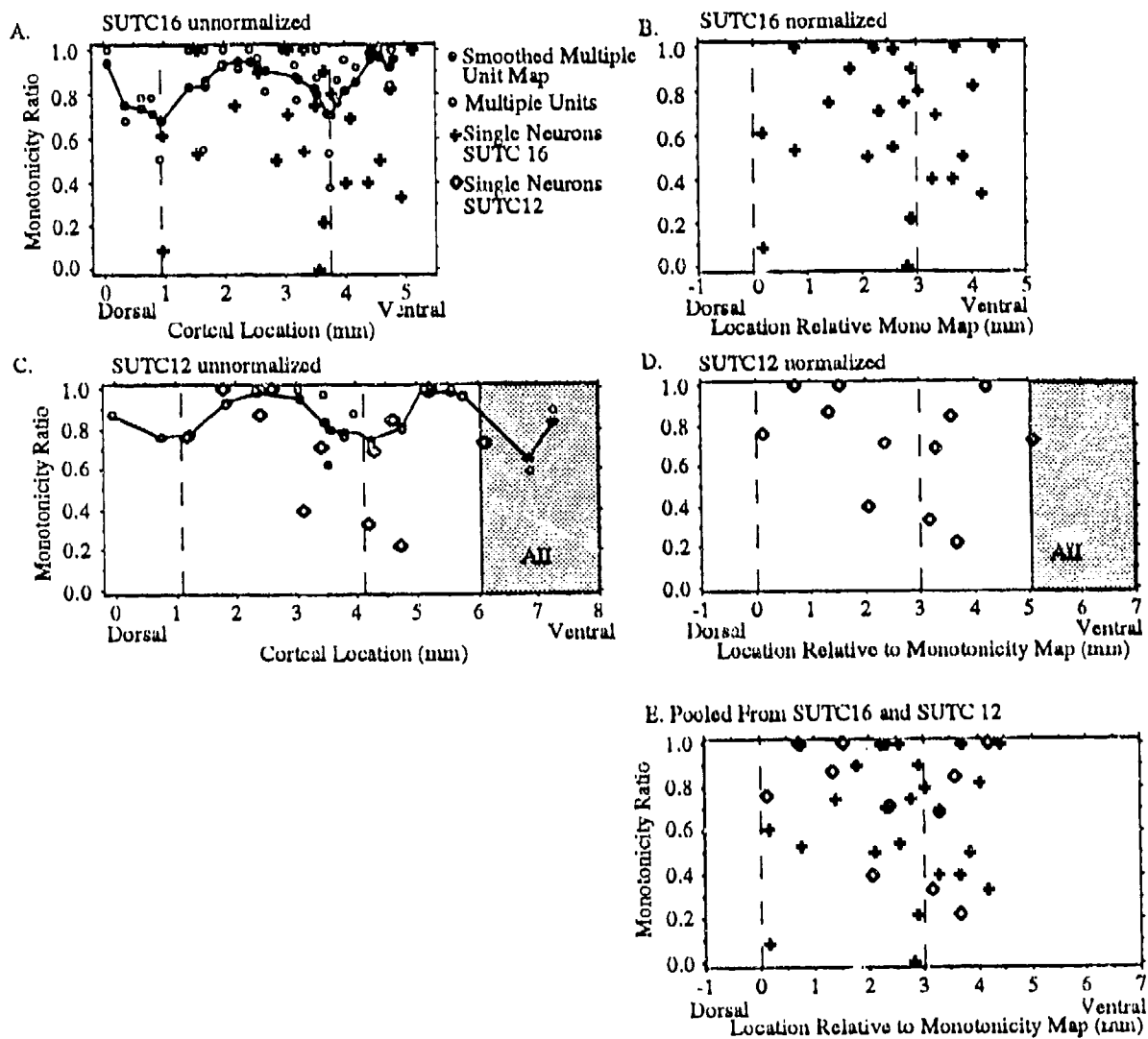
Fig 5
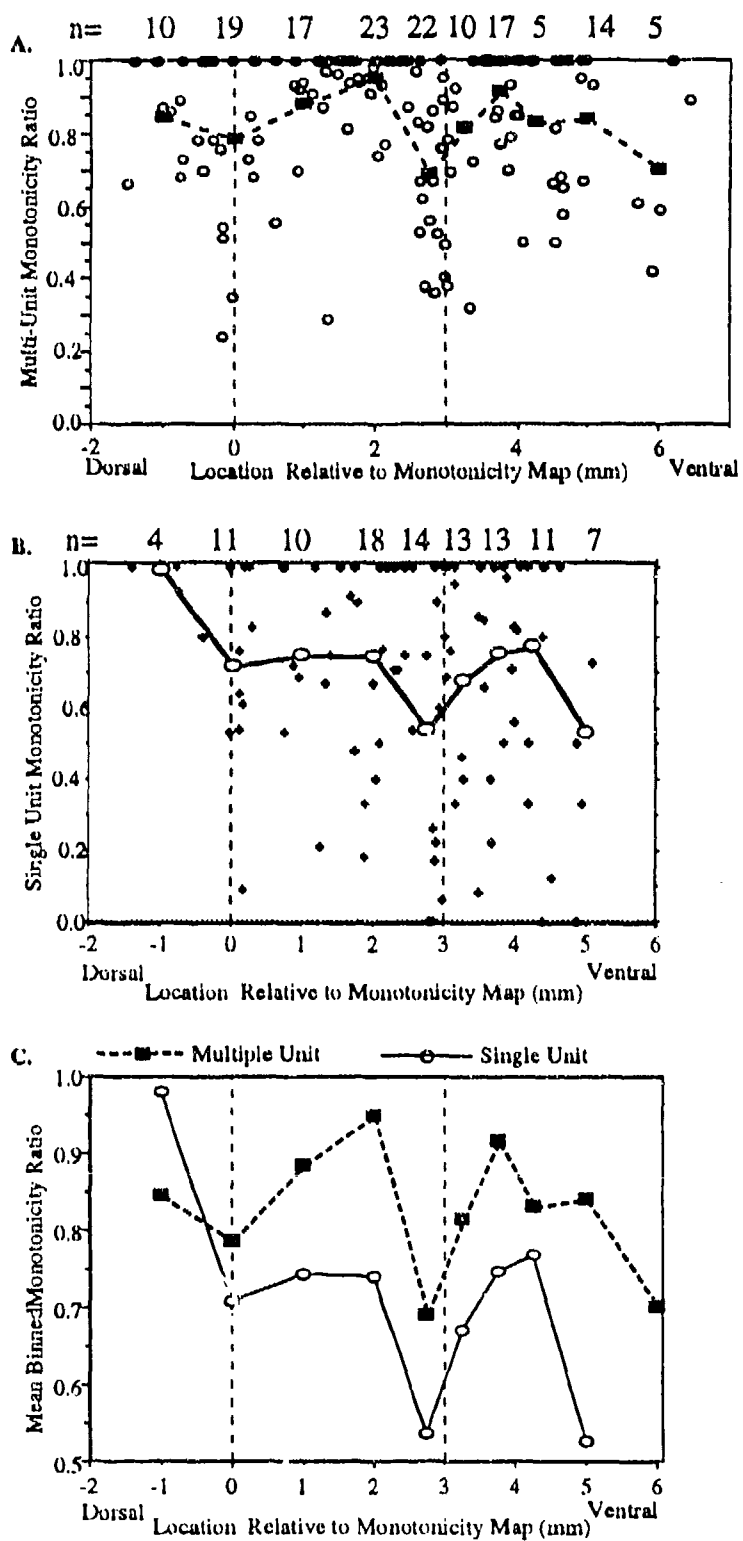
Figure 6

Strongly Non-monotonic (% Reduction < 50)

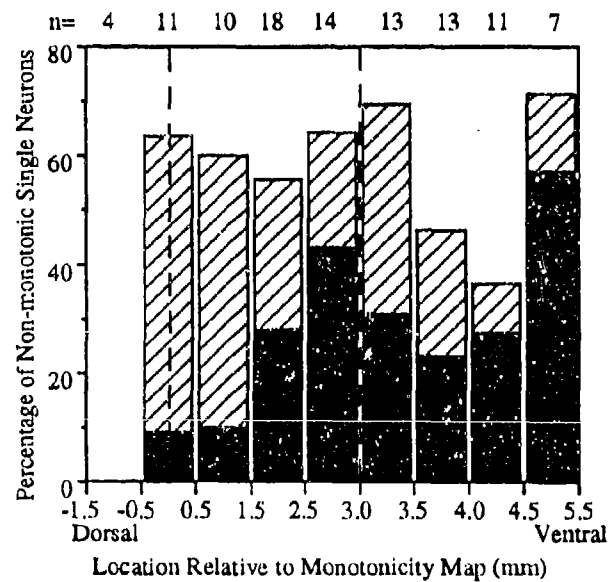Intermediately Non-monotonic (% Reduction Between 51 and 80)
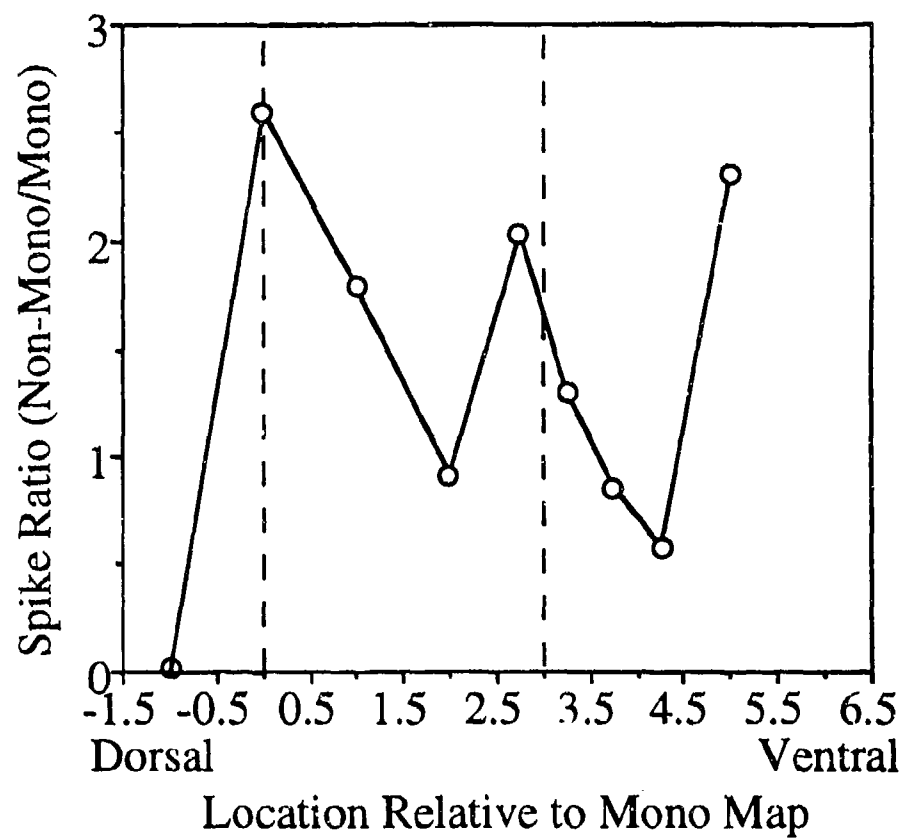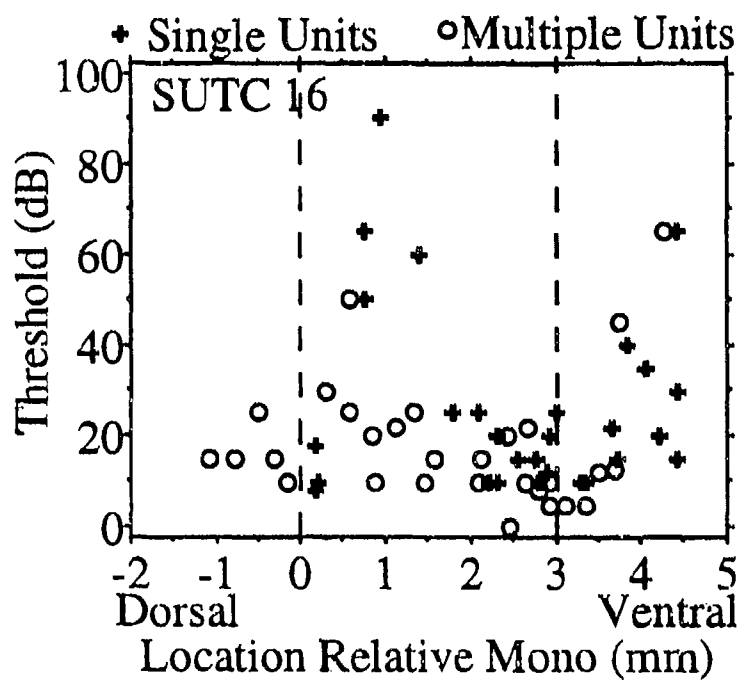
A.

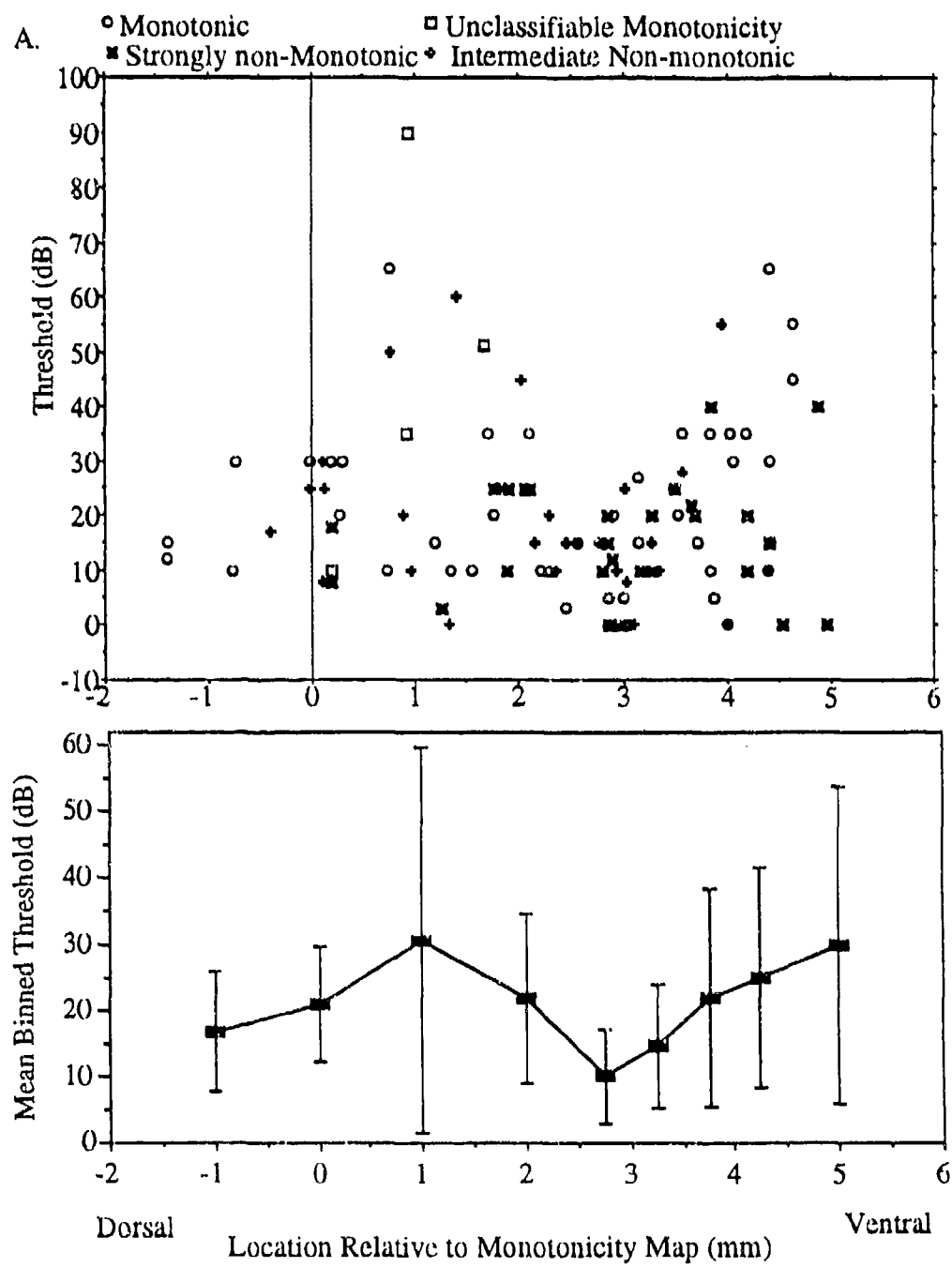n= 10 19 17 23 22 10 17 5 14 5

Figure 7
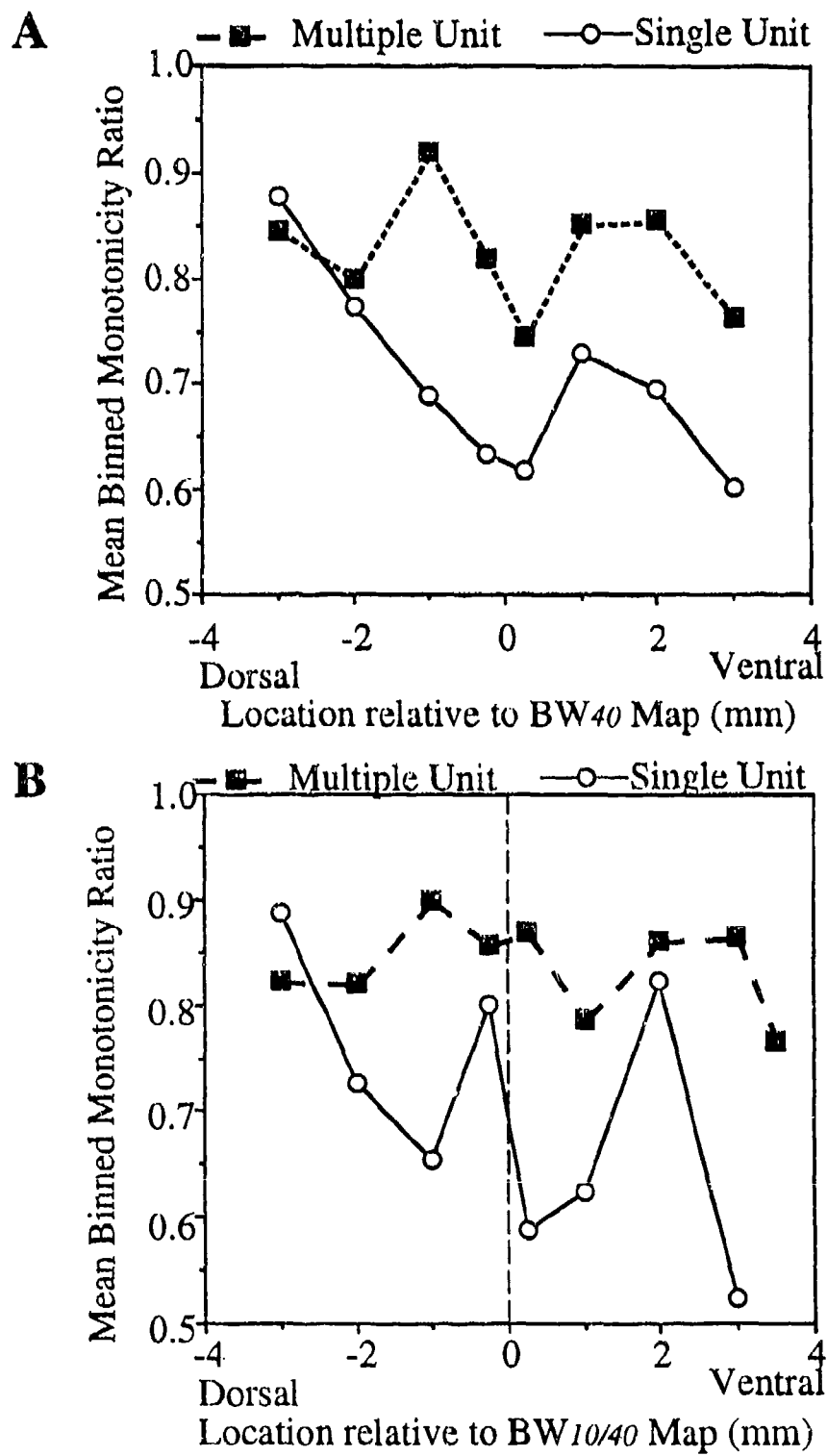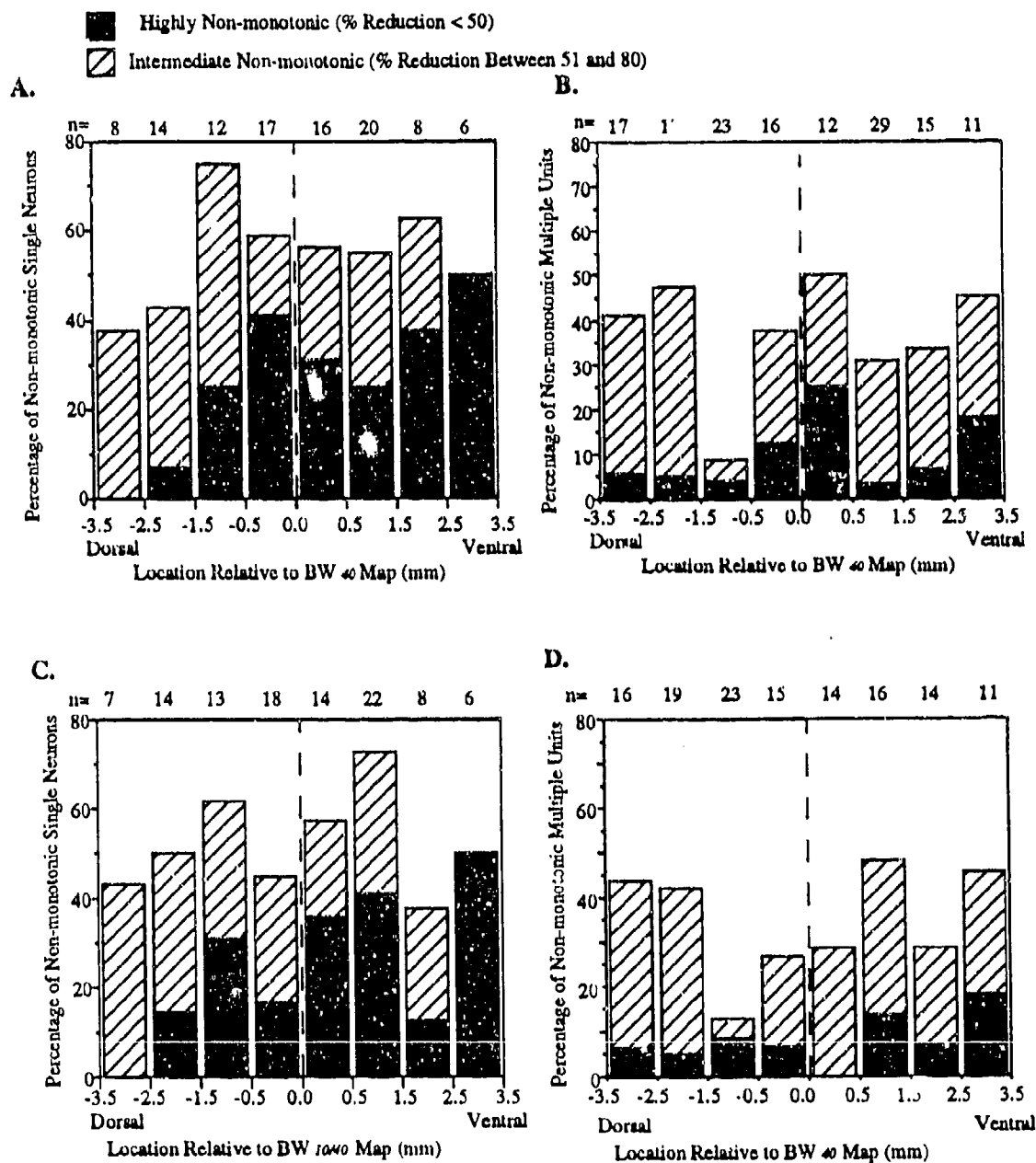
Figure 8

Figure 9

Figure 10:
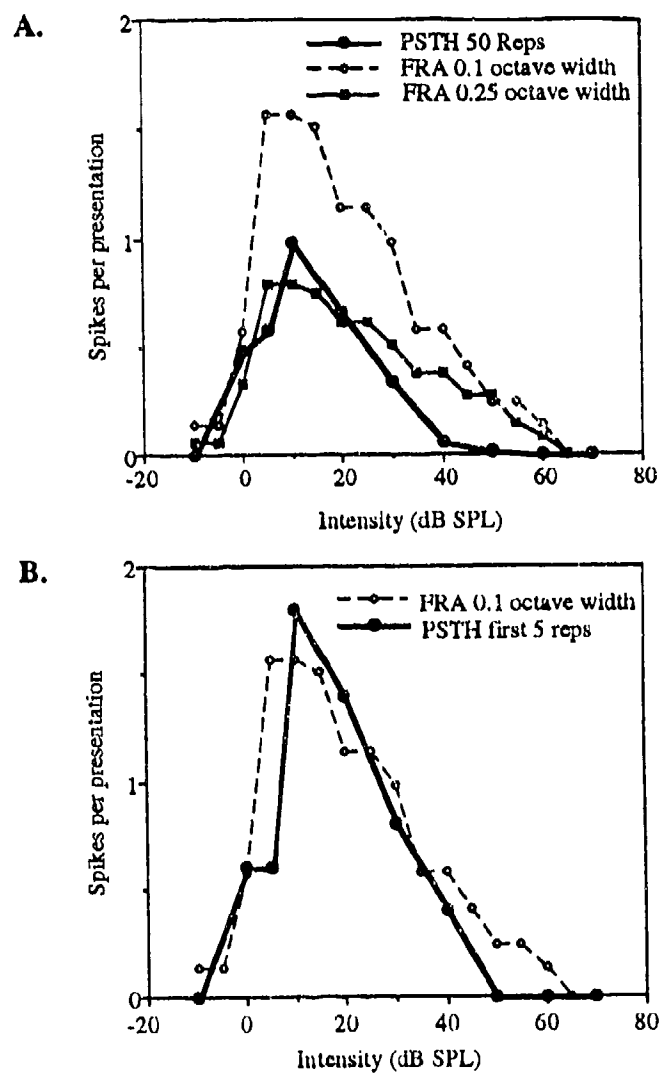
Figure 11:

Figure 12:

Figure 13

AUDITORY DISCRIMINATION LEARNING CHANGES NEUROPHYSIOLOGICAL RESPONSE OF CORTEX

M. Diane Keeling [a], Katharina Kruger [b], Barbara M. Calhoun, & Christoph E. Schreiner, Coleman Memorial Laboratory, W.M. Keck Center for Integrative Neuroscience, University of California, San Francisco; San Francisco, CA 94143-0732.

[a] now at Psychology Dept., Box 5050, University of New Brunswick, Saint John, Saint John, N.B. E2K 2E2, Canada.

[b] now at Institut fur Neuroinformatik, Ruhr - Universitat Bochum ND 04, Postfach 102148, D-44780 Bochum, Germany.

Abstract

We have previously reported the results of behavioral training of three cats which performed an auditory discrimination task involving complex stimuli that were created to simulate vowels - their frequency spectrum containing various peaks (vowel formants) located at various spacings. Their performance improved gradually and steadily to a final best difference threshold (Keeling et al, 1994). In this paper, we describe how the behavioral thresholds correlated with the electrophysiologically-measured Q-40 values of overall neural responses, indicating that a sharpening in tuning was related to the degree of proficiency in the auditory discrimination. We also report electrophysiological results from three cats trained with a slightly different stimulus, from whom behavioral thresholds could not be obtained, but whose cortical representation was changed nonetheless. The data imply that language acquisition involves dynamic cortical reorganization of speech feature representations and that human linguistic capabilities evolved from basic mammalian auditory processing capacities.

Speech perception is one of the most elaborate feats performed by the human brain, and while a number of models of speech perception have been posited and electrophysiological recording of cortical neuronal response to various acoustic signals has been performed, only speculative theories about the general physiological processes underlying speech perception have so far been possible. Consider the processing of vowels.

Vowel formants

Vowels are characterized by their formants - frequencies which are elevated in amplitude relative to other frequencies and which are produced by the natural resonances of the vocal tract activated during production of the vowel (indicated as F1, F2, F3, etc.). Formant values show considerable variability across speakers, however, and even overlap for some vowels (Peterson & Barney, 1952). The variability across speakers is due to physical differences in vocal tract size but also to a speaker's past experience, mainly his/her particular dialectical background, but the values may vary within context and even across repetitions as well.

In light of this variability it is amazing that listeners are so proficient at identifying vowels. The time course of a dialogue between two speakers is usually quite rapid and only rarely marked by mislabeling and misunderstanding of words. If there is no constant code between absolute formant frequency values and a psychological linguistic percept, how are vowels perceived?

Miller (1989) has proposed what he calls an "auditory-perceptual theory of phonetic recognition". The theory is descended from the formant-ratio theory of vowel quality, originally proposed by Lloyd (1890), who stated that vowel quality depends on the intervals between the formants, not their absolute values. The theory relies heavily on the

logarithmic frequency scale and Miller gives a detailed description of the effectiveness of various scales (e.g. mel, Bark, Koenig) in clustering vowel data (i.e. plotting formant values against each other as Peterson & Barney (1952) had done). The best grouping was observed by using the logs of ratios of formant center frequencies measured either in hertz or mels (the unit used for subjective pitch). Miller further justifies the use of a log frequency scale by citing Weber's law (namely that as stimuli are increased by multiplication, sensations increase by addition) and the octave basis of musical scales. The theory makes use of formant ratios and a "sensory reference" which is based on the speaker's average fundamental frequency. Sensory and perceptual "paths" describe changes in spectral patterns of the formants during the course of an utterance. "It is proposed that a segmentation mechanism based on the dynamics of these spectral patterns causes perceptual target zones to issue neural symbols or category codes that correspond to the vowel sounds" (p.2129).

Chistovich (1985) has reported a difference between stimuli with closely spaced formants and those with widely spaced formants when amplitude relations between the two formants are changed. Closely spaced formants were considered to be those with a distance between them of less than 3 - 3.5 bark (roughly equivalent to one octave). Their perception was closely related to the one-formant stimulus phonetically most similar to the two-formant stimulus, implying a loss of ability to discriminate the two peaks within this critical distance. She suggested therefore that spatial integration of spectral information may be an important factor in the processing of vowels. Chistovich speculates that "there exists a set of detectors tuned to specific spectral configurations...nearly simultaneous occurrence of two peaks in the dynamic spectrum is needed to excite the

detectors tuned to two-formant shapes" (p.803). She laments the absence of neurophysiological data concerning the response of neurons in AI of auditory cortex to stimuli with complex spectrum shapes.

The ratios of the formants may be an important cue. We calculated the ratios of the first and second, and second and third formants; since frequency is represented on a logarithmic scale along the basilar membrane within the cochlea, we computed the ratios of the formants in units per octave (Miller,1989) and hypothesized that the perception of variations in the position of spectral peaks is dependent on the spacing of the peaks.
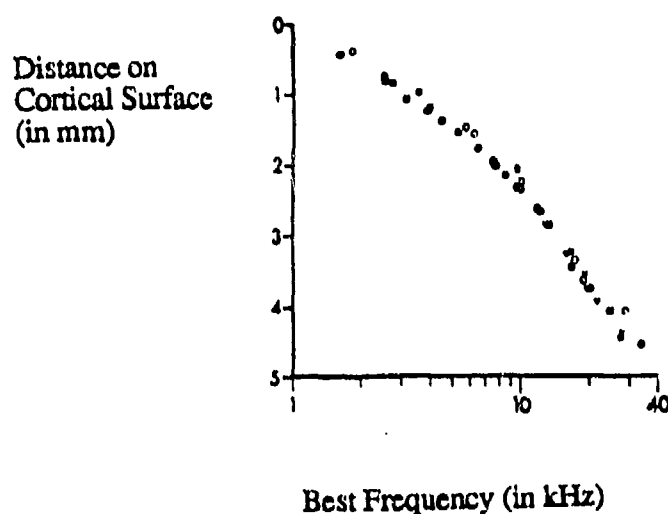
Auditory Cortex

The auditory cortex in the human and in many nonhuman primates is located primarily on the superior temporal gyrus and is buried for the most part within the Sylvian fissure (Lass, 1988). In the cat, the auditory cortex represents a large area mostly on the lateral surface of the brain. The following descriptions will refer to the cat.

On the basis of constant cellular characteristics seen with Nissl stain, Rose (1949) defined primary auditory cortex (termed AI), secondary cortex (AII), and an auditory area on the posterior ectosylvian gyrus (Ep). Primary auditory cortex was described as cytoarchitecturally similar to other primary sensory cortex, with six layers and a high density of pyramidal and granule cells in layers II, III and IV. The ventral division of the medial geniculate body projects to primary auditory cortex, the afferent fibers ending mainly in layer IV (Winer, Diamond, & Raczkowski, 1977).

Using surface evoked potential recordings, Woolsey and Walzl (1942) first reported a tonotopic organization in auditory cortex. In 1975,

Merzenich, Knight, & Roth refined this description employing microelectrode mapping techniques. They reported that "any given frequency band (or sector of the cochlear partition) is represented across a belt of cortex of nearly constant width that runs on a nearly straight axis across AI "(p.247) - see figure 1. Also, best frequency was constant within vertical penetrations into AI in the active middle and deep cortical layers, i.e. the cells appear to be organized in columns. Further, "there is an orderly representation of the cochlea within the field rostral to AI, with a reversal in best frequencies across its border with AI". Finally, "physiological definitions of AI boundaries are consistent with their cytoarchitectonic definition" (ibid).

Further mapping studies (Knight, 1977; Merzenich, Roth, Knight & Colwell, 1977; and Reale & Imig, 1977) have shown that there are at least three other cochleotopically organized cortical fields in cats, besides AI, namely, the anterior auditory field (AAF), which borders AI rostrally, the posterior auditory field (PAF), and the ventro-posterior field (VPAF) - see figure 2, from Merzenich, Andersen, & Middlebrooks (1979).

Figure 1. An example of a cortical map showing best frequency as a function of distance across the cortical surface (from Merzenich et al., 1975).

AUDITORY CORTICAL FIELDS
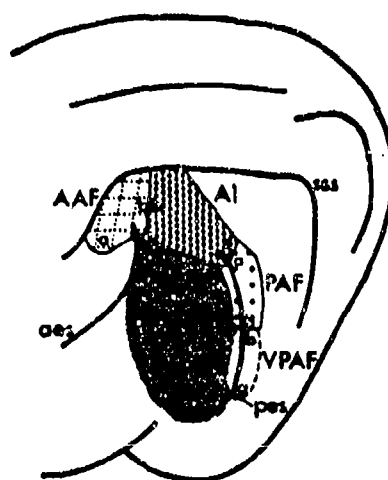
Figure 2. Schematic representation of the basic organization of auditory cortical fields in the cat. AI - primary auditory field; AAF - anterior auditory field; PAF - posterior auditory field; VPAF - ventroposterior auditory field. The region of representation of the cochlear apex (a) and cochlear base (b) are indicated for all four of these cochleotopically organized fields. (from Merzenich et al., 1979).

Another organizing feature of the auditory cortex is the existence of binaural bands roughly orthogonal to the iso-frequency contours (Imig & Adrian, 1978; Imig & Brugge, 1978; Middlebrooks, Dykes, & Merzenich, 1980). In alternating bands, "neurons are predominantly 'excitatory-inhibitory' (driven by contralateral stimulation with driven responses inhibited by ipsilateral stimulation); or, in the second set of bands, 'excitatory-excitatory' (driven by both contralateral and ipsilateral stimulation; or driven by one or the other ear, with the response strongly facilitated by stimulation of the other)" (Merzenich et al., 1979, p.65). It has also been shown that discrete frequency-band-specific lesions in unilateral primary auditory cortex produced profound deficits of sound localization ability in the contralateral hemifield (Jenkins & Merzenich, 1984).

There are however, some features of signals other than frequency and laterality which appear to be represented in a systematic way within auditory cortex. For example, sharpness of tuning is distributed along the iso-frequency domain of AI. Schreiner & Cynader (1984) have demonstrated low Q-10 dB values in area AII. Moving dorsally, the Q-10 dB values increased gradually, reaching a maximum about 2 mm dorsal to the AI/AII border. From the central portion of AI toward the Supra-Sylvian sulcus, the Q-10 dB values gradually decreased, although they remained higher there than in AII. This gradient in sharpness of tuning along the dorso-ventral extent of AI was confirmed by Schreiner & Mendelson (1990), the most dorsal and most ventral regions of AI being the most broadly tuned, with sharpest tuning in the central area. Amplitude modulated signals also appear to be represented somewhat systematically within the anterior auditory cortical area (Schreiner &

Urbas, 1986, 1988). Also, non-monotonicity of rate/level functions, i.e. the degree of change of firing rate as a function of level increase, has been investigated, and it appears that primary auditory cortex contains a topographic representation of intensity information in the iso-frequency domain (Schreiner, Mendelson, & Sutter, 1992; Heil & Scheich, 1991). Finally the direction and velocity of logarithmic frequency sweeps are systematically represented in AI (Mendelson, Schreiner, Grasse, & Sutter, 1988; Mendelson & Grasse, 1992). The most dorsal region of AI showed a preference for sweeps from low to high frequencies; moving ventrally, the preferred sweep direction gradually changed to sweeps from high to low frequencies, while most of the ventral portion of AI showed no specific sweep direction selectivity. FM-speed selective responses also displayed a systematic spatial organization, again progressing as a function of the dorso-ventral position within AI.

Suga (1984) has eloquently demonstrated the systematic representation of different types of biosonar information in the auditory cortex of the mustached bat. He shows how "complex acoustic signals are processed by specialized neurons that are tuned to particular information-bearing parameters (IBPs) or combinations of IBPs" (p.315). Furthermore, "the biologically more important values of an IBP are overrepresented by the large number of IBP filters tuned to them in order to achieve higher resolution" (ibid). Thus the highly refined localization ability of this animal has a corresponding highly refined cortical topographical representation. He has hypothesized (1988) that speech recognition is similarly based upon a spatio-temporal pattern of neural activity occurring in various cortical areas.

Wollberg & Newman (1972; Newman & Wollberg, 1973) looked at the responses of cells in the superior temporal gyrus (STG) of awake squirrel monkeys to a variety of different species-specific vocalizations. Eighty-nine percent of the cells sampled responded to over half of the vocalizations, suggesting that "many neurons in the STG do not select between different classes of vocalizations according to presence or absence of simple acoustic features" (Newman & Wollberg, 1973, p.287), and that vocalizations are encoded in a highly complex way.

Langner, Bonke, & Scheich (1981) explored neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. Field L is a layered and tonotopically organized primary auditory projection area in the bird. Only a minority of units responded selectively to one particular vowel; many units responded to several vowels. It was therefore speculated that vowel recognition may be based on populations of simultaneously activated units.

Steinschneider et al. (1990) recorded activity in AI in an awake monkey to 3 consonant-vowel syllables, and to the syllables' isolated formants and formant pairs. They reported that response features were related to the tonotopic organization of AI and that formant interactions seem to modulate the response to whole syllables.

Shamma, Fleshman, Wiser, and Versnel (1993) studied responses in primary auditory cortex of the ferret to direction of frequency-modulated tone sweeps and to spectrally shaped noise. The excitatory and inhibitory portions of the response area were described in terms of the asymmetry of the excitation and inhibition around the best frequency and were shown to correlate with the preferred direction of FM sweeps and with spectral shapes. The authors concluded that "cortical responses encode the locally

averaged gradient of the acoustic spectrum by their differential distribution along the isofrequency planes. This enhances the representation of such features as the symmetry of spectral peaks and edges and the spectral envelope" (p.367). They suggest a possible analogy in AI with spatial frequency "channels" in the primary visual cortex.

Schreiner, Calhoun, and Keeling (1993) presented ripple stimuli to cortical neurons in the cat. The ripple spectra that were generated had the following characteristics: the carrier consisted of a harmonic series (F0 ranging from 50 to 200 Hz) with a 6 dB/octave decline of the component amplitudes (120 to 255 components); the bandwidth of the stimulus was 3 octaves; the spectral envelope of the signal was represented by a sinusoid on a logarithmically scaled frequency axis, the frequency of the envelope sinusoid was referred to as ripple density (ripples/octave); the modulation depth of the envelope (ripple depth) was linear on a dB scale. The geometrical center of the band-limited signal was always at a maximum of the sinusoidal spectral envelope. The center was positioned at the characteristic frequency (cf) of each cortical neuron and the ripple density or the frequency distance between spectral peaks was systematically varied. The resulting 'ripple transfer function' was reconstructed for different modulation depths and overall intensities. For the majority of neurons, the ripple transfer function was a bandpass and a 'best' ripple density could be defined. The remaining transfer functions appeared to have a lowpass characteristic, at least for those ripple densities used (0.3 to 8 ripples/octave). Best ripple densities ranged from 0.6 to 4 ripples/octave with a mean between 1 and 2 ripples/octave. Spatial mapping of responses to ripple spectra along the isofrequency domain of AI revealed a systematic shift of the best ripple density in multiple unit responses from central AI to

dorsal AI. The shift paralleled the previously described variation of integrated excitatory bandwidth with sharply tuned locations and high best ripple densities near the dorso-ventral center and broader tuning curves with lower best ripple densities toward the dorsal end of AI. Ventral AI showed a less systematic distribution of ripple densities. The hypothesis of a systematic 'spatial frequency' or 'spectral envelope frequency' representation in AI, oriented orthogonal to the frequency axis, was supported.

Thus, one sees at the primary auditory cortical level a rudimentary representation of spectral and temporal features of speech or speech-like signals. What is <u>not</u> seen is selectivity to particular calls, or syllables, or vowels. Rather neural responses seem to be evoked by discrete stimulus components. What is supported is the idea that percepts are created by patterns of neural firing across the length and breadth and probably depth of the sensory cortical area.

Plasticity of Auditory Cortex

Merzenich and colleagues have provided evidence that, in adult mammalian somatosensory cortex, the topographic map of the body surface is not static, and that receptive fields at particular cortical sites can change in size and location throughout adult life; the cortical locus at which a given skin surface is represented can shift several hundreds of microns across the cortex (Clark, Allard, Jenkins, & Merzenich, 1988; Jenkins, Merzenich, Ochs, Allard, & Guic-Robles, 1990; Allard, Clark, Jenkins, & Merzenich, 1991). They have suggested that "inputs are selected on the basis of temporal correlation" (Clark et al, 1988, p.444), i.e. mechanisms which are correlated in time may underlie cortical representations and

their modification. Such changes also appear to underlie the recovery seen following damage to the brain (Jenkins & Merzenich, 1987). It has also been shown that extended use of a monkey digit, in a discriminative task for which reward is contingent upon behavior, induces reorganization/modification of somatosensory cortical representation (Recanzone, Jenkins, Hradek, & Merzenich, 1992; Recanzone, Merzenich, Jenkins, Grajski, & Dinse, 1992).

In auditory cortex, Robertson and Irvine (1989) have shown that frequency organization in guinea pig reorganizes following partial unilateral deafness. In cats, the cochleotopic representation in primary auditory cortex (AI) was extensively reorganized following neonatal, bilateral high frequency cochlear damage and hearing loss (Harrison, Nagasawa, Smith, Staton, & Mount, 1991). In owl monkeys, Recanzone, Schreiner, and Merzenich (1993) observed plasticity in the frequency representation of primary auditory cortex following discrimination training. An increase in the cortical area of representation of those frequencies which were behaviorally relevant to the monkeys was observed. The sharpness of tuning was increased, as was the latency of the cortical responses.

Thus, auditory cortical representation of acoustic stimuli is not static, but changes with changing relevance of the stimuli. This dynamic adjustment of cortical response appears to underlie the learning process, as well as recovery following injury.

Rationale

Thus, the perception of vowels involves the processing of peaks of spectral energy, the frequencies of which are located at various distances or

spacings. It was decided to create a vowel-like "ripple" stimulus, which could be easily manipulated to allow the presentation of various parameters of the basic stimulus. The main parameter of interest was to be the peak spacing or ripple density. Holding the ripple density constant, in each test condition, the phase of the envelope of the stimulus was to be shifted so that the locations of the peaks changed. The precision with which the peak shifts could be determined, and the variation of this precision across several ripple densities was expected to shed light upon the importance of small changes in formant location for vowel processing. An operant conditioning paradigm was employed to obtain a threshold for discriminability of phase shifts of a ripple stimulus in cats.

It was postulated that at the primary auditory cortical level, a vowel-like ripple spectrum would be represented on the basis of the frequency of the dominant peaks in conjunction with a representation of formant ratio or peak spacing as processed by the filter function (sharpness of tuning) of the driven neurons. It was further postulated that, as for plastic changes seen in primate cortex consequent to behavioral training, the cortical representation of behavioral stimuli would be different, reflecting more efficient processing of these relevant stimuli, in the cortex of trained animals as compared to that in untrained animals.

METHODS

Electrophysiological recording from primary auditory cortex was performed in six trained cats, as well as in several normal un-trained cats. The cats were injected intra-muscularly with acetylpromazine maleate (.09 mg/Kg) and ketamine hydrochloride (10 mg/kg). When a cat was sufficiently sedated, it was shaved along the inside of the two forelimbs (for IV cannulation), along the throat area (for insertion of tracheal tube), and over the posterior and lateral area of the head (for temporal cortex exposure). Venous cannulation was performed; anesthesia was induced with an initial dose of pentobarbital sodium (30 mg/kg). Animals were maintained at a surgical level of anesthesia with a continuous infusion of pentobarbital sodium (2 mg $\cdot$ Kg^{-1} $\cdot$ h^{-1}) in lactated Ringer solution (infusion volume: 3.5 ml/h) and, if necessary, with supplementary intravenous injections of pentobarbital sodium. The state of anesthesia was monitored and the rate of infusion was adjusted to maintain an a-reflexive state. The cats were also given dexamethasone sodium phosphate (0.14 mg/kg im) to prevent brain edema and atropine sulfate (1 mg im) to reduce salivation. A tracheotomy was performed and a tracheal tube inserted to ease breathing and to reduce breathing noises. The temperature of the animal was monitored with a rectal temperature probe and maintained at 37.5 C with the use of a heating water blanket with feedback control. The EKG was continuously displayed on an oscilloscope and amplified through a loudspeaker.

The cat's head was placed in a standard rigid mouth-bar headholder, an incision made in the scalp down the center from forehead to occiput, the right temporalis muscle retracted, and the right cortical surface (right

side for convenience due to stimulation and recording set-up) exposed
by drilling through the skull in the region of the middle ectosylvian gyrus.
The dura overlying the middle ectosylvian gyrus was removed, the exposed
cortex covered with silicone oil, and a video image of the surface
vasculature was obtained with a CCD camera, an image capture board
(Data Translation DT2255) and capture software (Image 1.4, NSCA).
Electrode penetration sites were marked on the image of the cortical
surface in a display program (Canvas, Deneba).

Experiments were conducted in a sound-shielded room (IAC).
Auditory stimuli were presented via calibrated headphones (STAX 54)
enclosed in small chambers that were connected to sound delivery tubes
placed into the acoustic meati (Sokolich, US Patent 4,251,686; 1981). The
sound delivery system was calibrated with a sound level meter (Bruel &
Kjaer) and a waveform analyser (General Radio 1521-B). Auditory
stimuli were digitally generated by a signal processing computer
(TMS32010) and converted into analog signal by a 16-bit digital-to-analog
converter running at a 60 kHz sampling rate. Additional attenuation was
provided by a pair of passive attenuators (Hewlett Packard).

For each recording site responses were firstly recorded to at least
675 different tone bursts. Tone bursts were presented in a pseudorandom
sequence of different frequency-level combinations selected from 15 level
values and 45 frequency values. From the responses to all stimuli,
frequency response areas (FRAs) were reconstructed. Steps between levels
were 5 dB, resulting in a sampled dynamic range of 75 dB. The frequency
range covered by the 45 frequency steps was geometrically centered
around the estimated cf of the recording site and covered 3 to 5 octaves,
depending on the estimated width of the FRA. Stimulus frequencies were

equidistant on a logarithmic frequency scale. Following the presentation of pure tones and the generation of the FRA, ripple stimuli with characteristics similar to the stimuli used in animal behavioral testing were created and presented: bandwidth was 3 octaves, fundamental frequency was approximately 69 Hz, ripple densities (10 to 14 values) of from 0.5 to 8 ripples/octave were presented, envelope phase shifts (10 values) of from 0 to 180 degrees were presented. See figure 3.
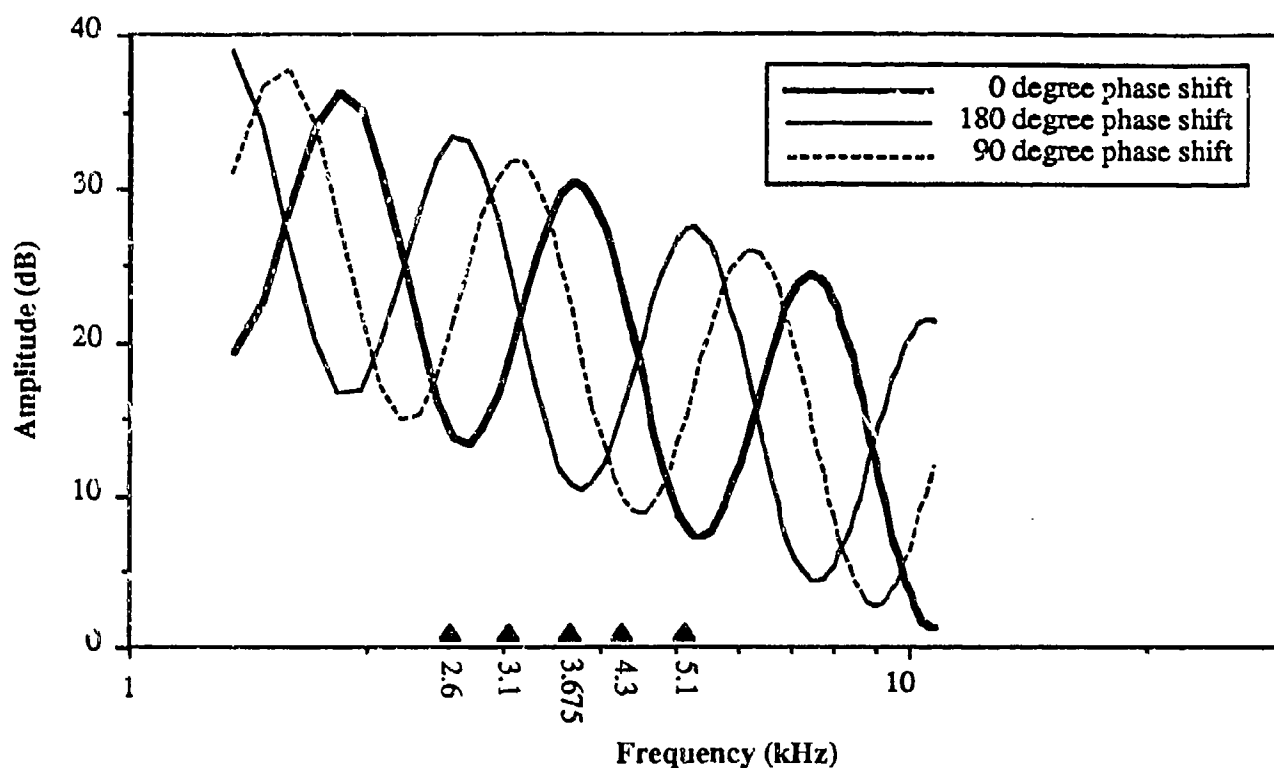
Figure 3.   Ripple Stimulus, showing 0 degree, 180 degree, and 90 degree phase shift. Corresponding frequency locations are shown: 3.675, 2.6, and 3.1 kHz.   Frequencies to the right represent phase shifts of negative 90 and negative 180 degrees, at locations of 4.3 and 5.1 kHz.   Ripple density = 1 ripple/octave.

Primary auditory cortex (AI) was identified using cf information, high frequencies (20 kHz) being represented anteriorly and low frequencies (2 - 4 kHz) posteriorly (Merzenich et al., 1975) and using sulcal/gyral landmarks. Multiple unit recording was performed using parylene coated tungsten electrodes with impedances of 1.0 to 2.0 MOhm at 1 kHz; in a few cases, single unit responses were recorded with glass-coated tungsten electrodes (3-4 MOhm at 1 kHz). The electrode was introduced into the cortex approximately orthogonal to the surface (as viewed through a Zeiss operating microscope); penetrations were parallel to each other. The electrode was advanced through the cortical layers to a depth of 600 - 1000 μm, corresponding to cortical layers III and IV, using a hydraulic microdrive (Kopf) remotely controlled by a stepping motor. Activity of small groups of neurons was amplified, band-pass filtered (1-10 kHz, 12 dB octave), and monitored on an oscilloscope and an audio monitor. The discriminator (BAK DIS-1) level was set to exclude evoked potentials and to accept events that resembled action potentials of an amplitude at least 50% above the background signal. The number of events per presentation and the arrival time of the first event after the onset of the tone bursts were recorded and stored in a computer (DEC 11/73). The recording window had a duration of 50 ms, corresponding to the stimulus duration. Poststimulus time histograms (PSTHs) were constructed in response to the ripple stimuli. Binwidth was 0.8 ms.

A general response profile of cells was assessed including spontaneous discharge rate, characteristic frequency (cf) (the stimulus frequency with the lowest sound-pressure level necessary to evoke neuronal activity), threshold (the lowest level evoking activity in the FRA), bandwidth approximately 10 dB above threshold (to compute Q-10, which

is cf divided by bandwidth 10 dB above threshold), bandwidth 40 dB above threshold (to compute Q-40, which is cf divided by bandwidth 40 dB above threshold), and in some cases binaural interaction class. Following electrophysiological testing animals were sacrificed with an overdose of pentobarbital sodium; bilateral thoracotomy was performed.

RESULTS

Several (eighteen) normal "untrained" cats provided data for this thesis. The recording protocol changed slightly over time as information about neural responses was gathered. Thus, not all untrained cat data are included in all analyses. Three cats, #218, 293, and 176 were behaviorally trained and tested on a stimulus with a ripple density of 1 ripple/octave; they underwent recording at the end of behavioral testing. Three other cats (#92-1767, CD184, and 92-1688) were behaviorally trained and tested on a stimulus with a ripple density of two ripples/octave. Two "exposed" cats, #122, and 40764, i.e. cats that had been placed on the training procedure but had been unable to learn the discrimination paradigm, also provided electrophysiological data.

Ripple Density Representation in Cortex

Ripple stimuli of various densities were presented during electrophysiological recording to both trained and untrained cats. In this section of the results, the spike counts recorded in the post-stimulus histograms were normalized, that is, for the responses to the ripple stimulus, the best response was designated as 100% and the smallest 0; all other responses were computed as a percentage between 0 and 100. In this way, variation in the responses to the ripple could be more easily observed, and the responses could be compared to one another, and across animals. The "neural response" was defined as the normalized spike count averaged across all penetrations, and described as a function of the ripple density of the stimulus. For electrophysiological response recording to different ripple densities, the ripple stimulus was centered on the cf of the multi-unit penetration, then various ripple densities (0.5 to 8 ripples/octave) were presented.

21

The electrophysiological results for cats trained with a ripple density of 1 ripple/octave will be discussed initially; untrained data will then be presented; the results from these trained and the untrained animals will then be compared; finally the data from all 6 trained animals will be compared with the results from untrained animals.

Figure 4a shows spike counts averaged across all penetrations as a function of ripple density for the first trained cat, #218. The total number of spikes, recorded over thirty repetitions of the ripple stimulus presentation, during the 30 ms following response onset, across all penetrations, for each ripple density, were added and averaged. The best response occurred to a ripple density of 1 ripple/octave. This function may be described as a cosine ripple transfer function (cosine, because it represents the stimulus presented only at a phase of 0 degrees; a complete description of the ripple transfer function also requires the response to the stimulus at phase 90 degrees). The bandwidth halfway between the peak and the lowest value within the function may be measured. In this case, the bandwidth 50% down from the peak is 2.33 ripples/octave.

Figure 4b illustrates the neural response as a function of ripple density for the second trained cat, #293. In this animal the best response was at a ripple density of 0.33 ripples/octave. The bandwidth 50% down from the peak was 1.0 ripple/octave.

Figure 4c illustrates the neural response as a function of ripple density for the third trained cat, #176. The best response was at a ripple density of 1 ripple/octave (0.33 ripple density is quite close; the peak at 4.66 was the neural response to the ripple density presented first in the series). The bandwidth 50% down from the peak was 1.66 ripples/octave.

Figure 4d illustrates the neural response as a function of ripple density averaged across these three trained cats. The best response was at a ripple density of 1 ripple/octave. The bandwidth 50% down from the peak was 1.75 ripples/octave.
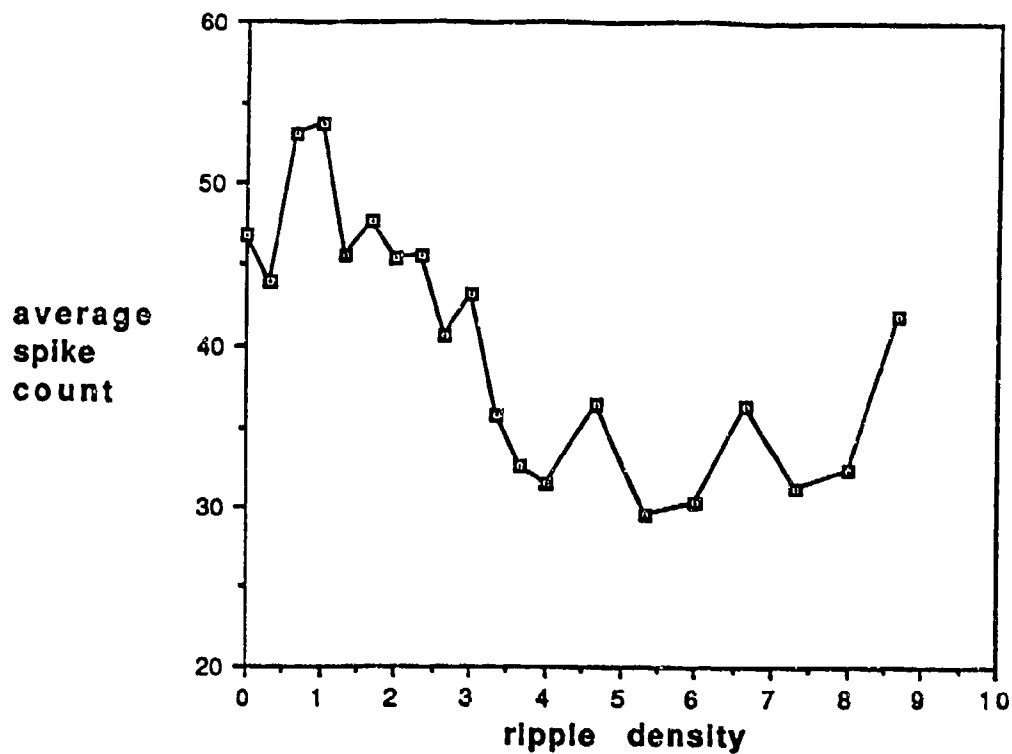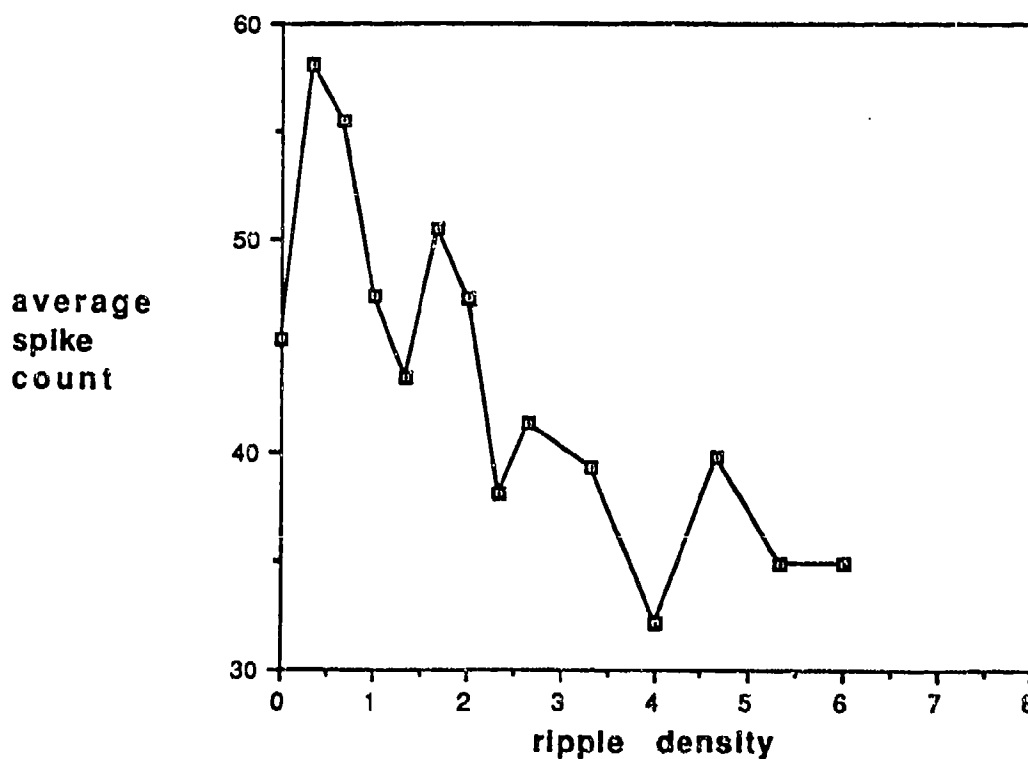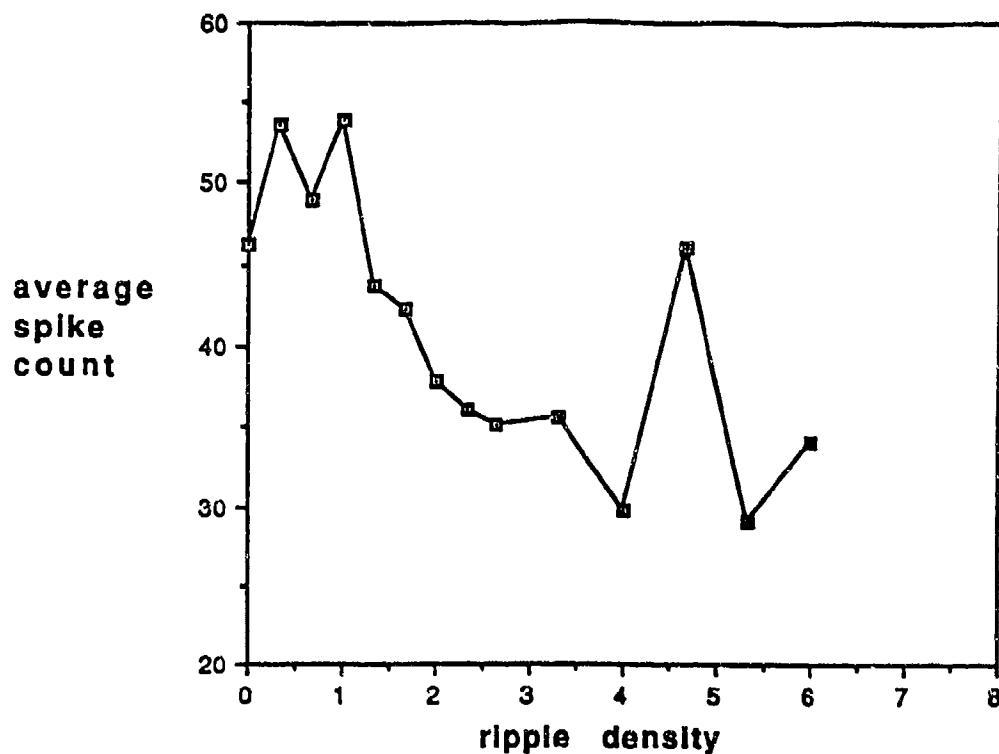
a) Cat #218

b) Cat #293

Figure 4 a) Neural response as a function of ripple density for Cat #218, and b) Neural response as a function of ripple density for Cat #293.
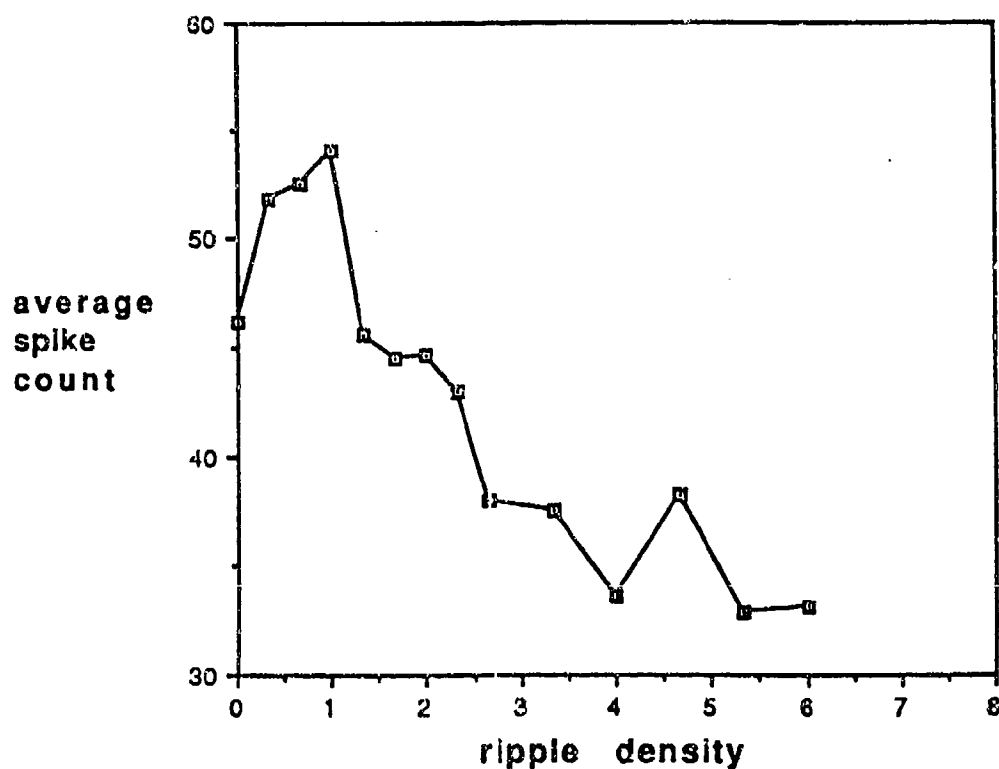
d) Average of three Trained cats

Figure 4 c) Neural Response as a function of ripple density for Cat #176, and d) Neural Response as a function of ripple density averaged across all three Trained Cats.

Figure 5 a,b,c,d,e, and f presents similar functions for five untrained cats (424, 188, 1681, 1673, and 92-1700) and for the average of the five untrained cats.

Table 1 presents the peak responses, bandwidths 50% down from the peak, number of microelectrode penetrations ('units'), and range of characteristic frequencies (cfs) included in the neural response for each of the three trained cats and for the five un-trained cats.

cat #	training status	r.d. peak response	bandwidth 50% down	number units	range of cfs (kHz)
218	Trained	1	2.33	45	2.3 to 9
293	Trained	0.33	1.2	77	2.3 to 11.8
176	Trained	1	1.66	50	2.7 to 11.7
681	Untrained	0.66	1.0	15	1.9 to 11.5
1673	Untrained	0.33	1.0	17	2.6 to 7.2
92-1700	Untrained	0.33	0.5	11	4.3 to 13.4
424	Untrained	0.66	1.33	18	1.5 to 7.9
188	Untrained	0.33	1.0	26	1.1 to 9.4

Table 1. Bandwidths at Best Ripple Response for Trained and Untrained Cats

Performing a Mann-Whitney U two-group un-paired comparison of the trained animals' bandwidths to those of the five untrained cats reported a significant difference (p=0.05). That is, the bandwidth 50% down from the best response to a particular ripple density was wider in trained cats than in untrained cats.
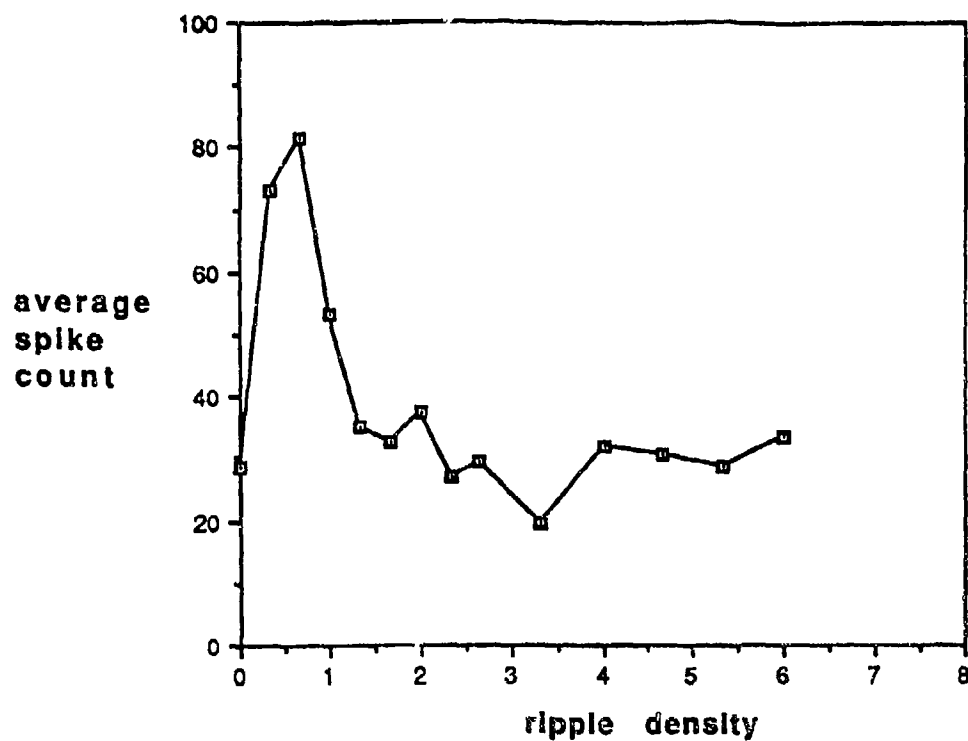
a) Cat #681

b) Cat #1673

Figure 5a) Neural response as a function of ripple density for Cat #681, and b) Neural response as a function of ripple density for Cat #1673.
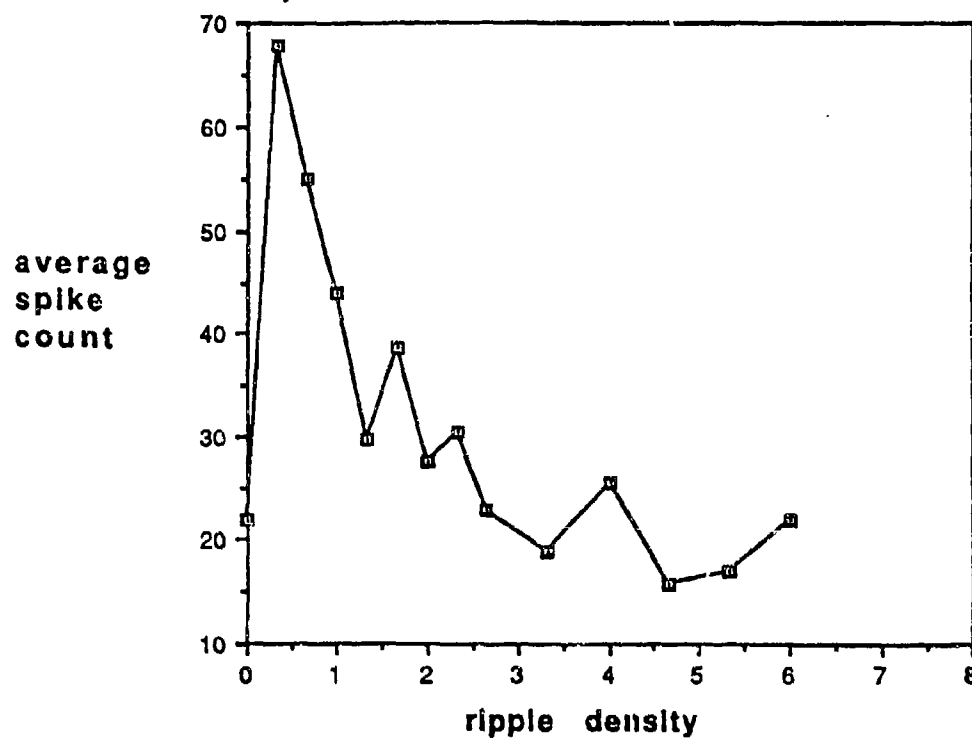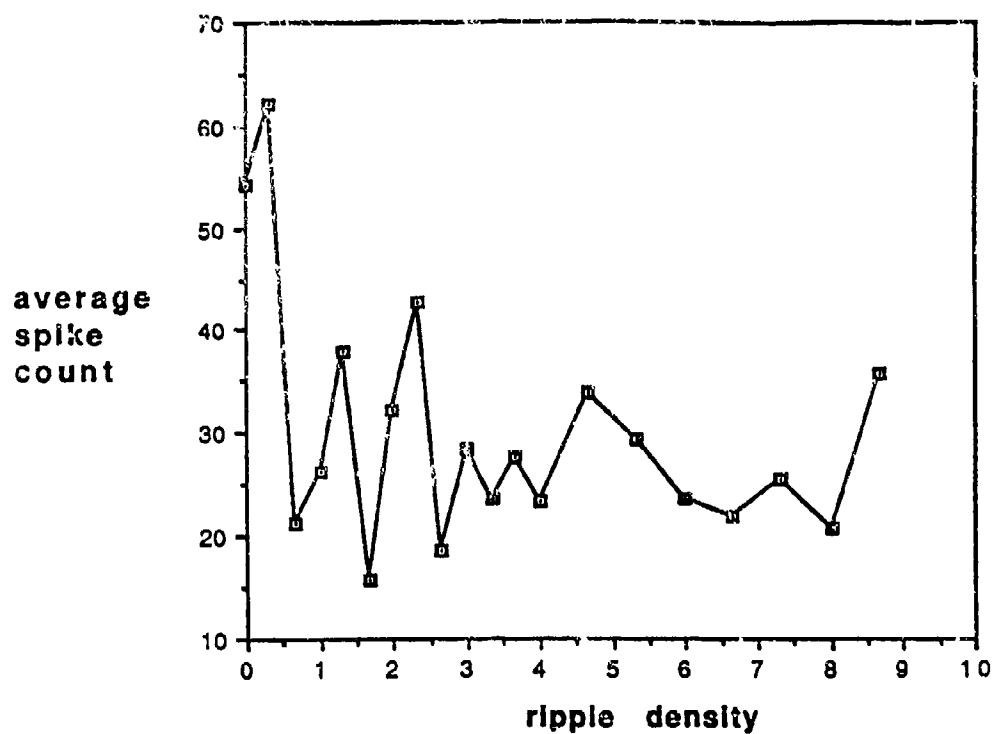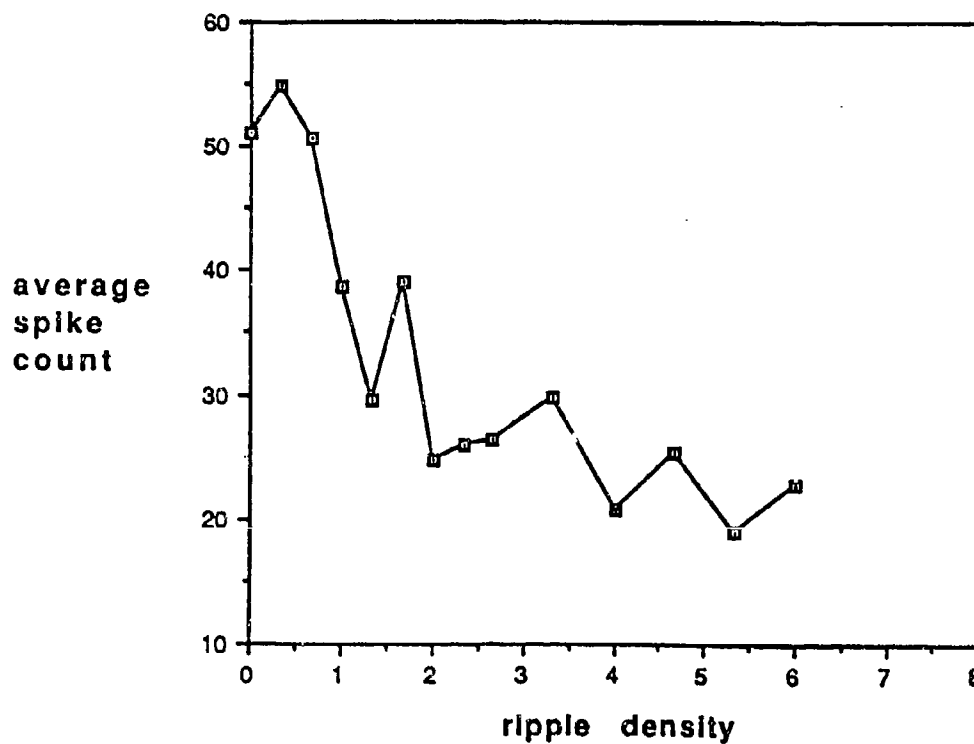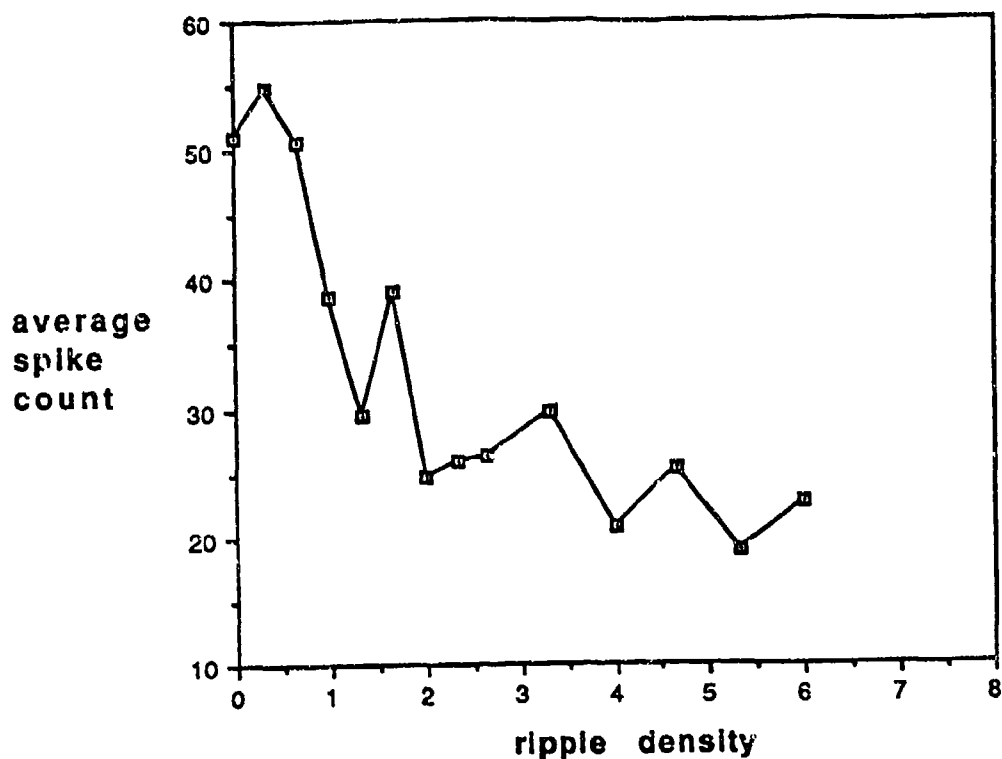
Figure 5 c) Neural response as a function of ripple density for Cat #92-1700, and d) Neural response as a function of ripple density for Cat #424.
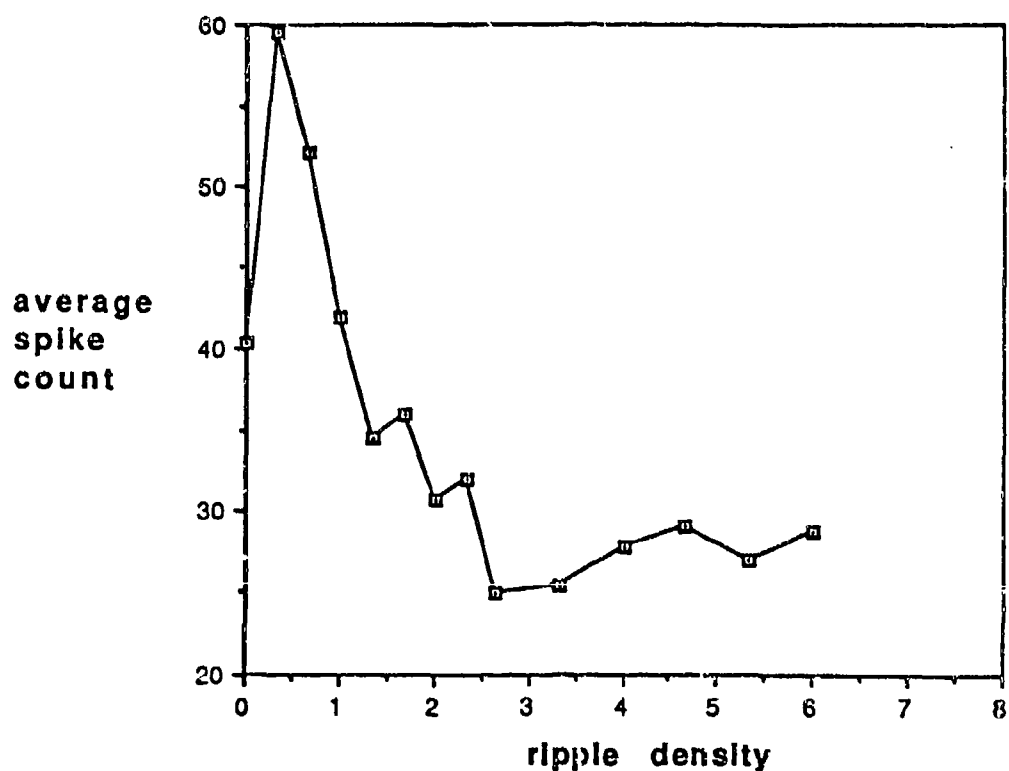
f) Average of 5 Untrained cats

Figure 5 e) Neural Response as a function of ripple density for Cat #188, and f) Neural Response as a function of ripple density averaged across five Untrained Cats.

Figure 6 presents a comparison of the neural response as a function of ripple density for the average of three trained cats and the average of five untrained cats. Performing a Wilcoxon Signed-Rank two-group paired comparison reported a difference significant at a level of $p=0.001$. Thus, the response to ripple densities was different in the trained cats as compared to the untrained cats. In trained cats, the best response was recorded in response to a higher ripple density than the best response in untrained cats. Overall, i.e., across all ripple densities presented, the responses were on average higher in the trained cats, except for the best response ripple density in untrained cats.

Thus, untrained cats showed the highest response to ripple densities of either 0.66 or 0.33 ripples/octave, with the mean best ripple response being at a ripple density of 0.33 ripples/octave. For three cats trained on a ripple density of 1 ripple/octave, the highest response was at either 1 or 0.33 ripples/octave, with the mean best ripple response being at a ripple density of 1 ripple/octave. The bandwidth of the ripple transfer function 50% down from the peak response was wider in these trained cats as compared to untrained cats.

Figure 7 illustrates the ripple transfer function for the cats trained on a ripple density of 1 ripple/octave, as above, as well as the three cats trained on a ripple density of 2 ripples/octave, compared to the ripple transfer function for the average response of 8 untrained cats.
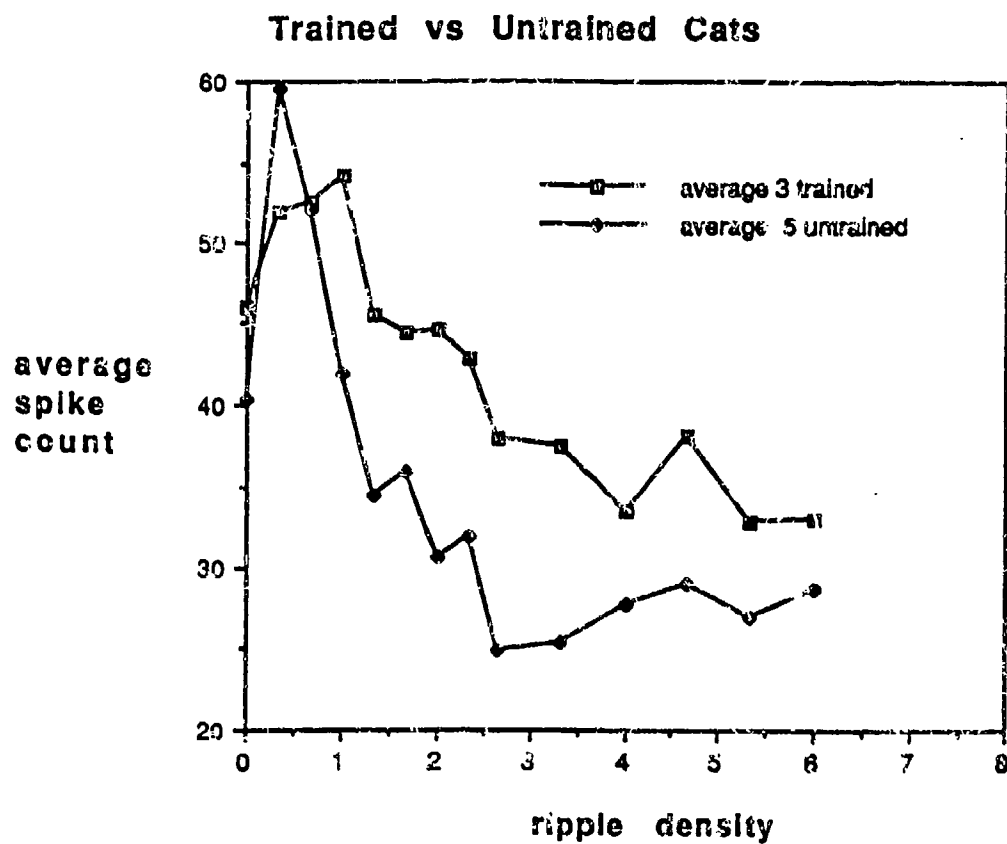
Figure 6. Comparison of Ripple Transfer Function for Trained Cats vs Untrained Cats
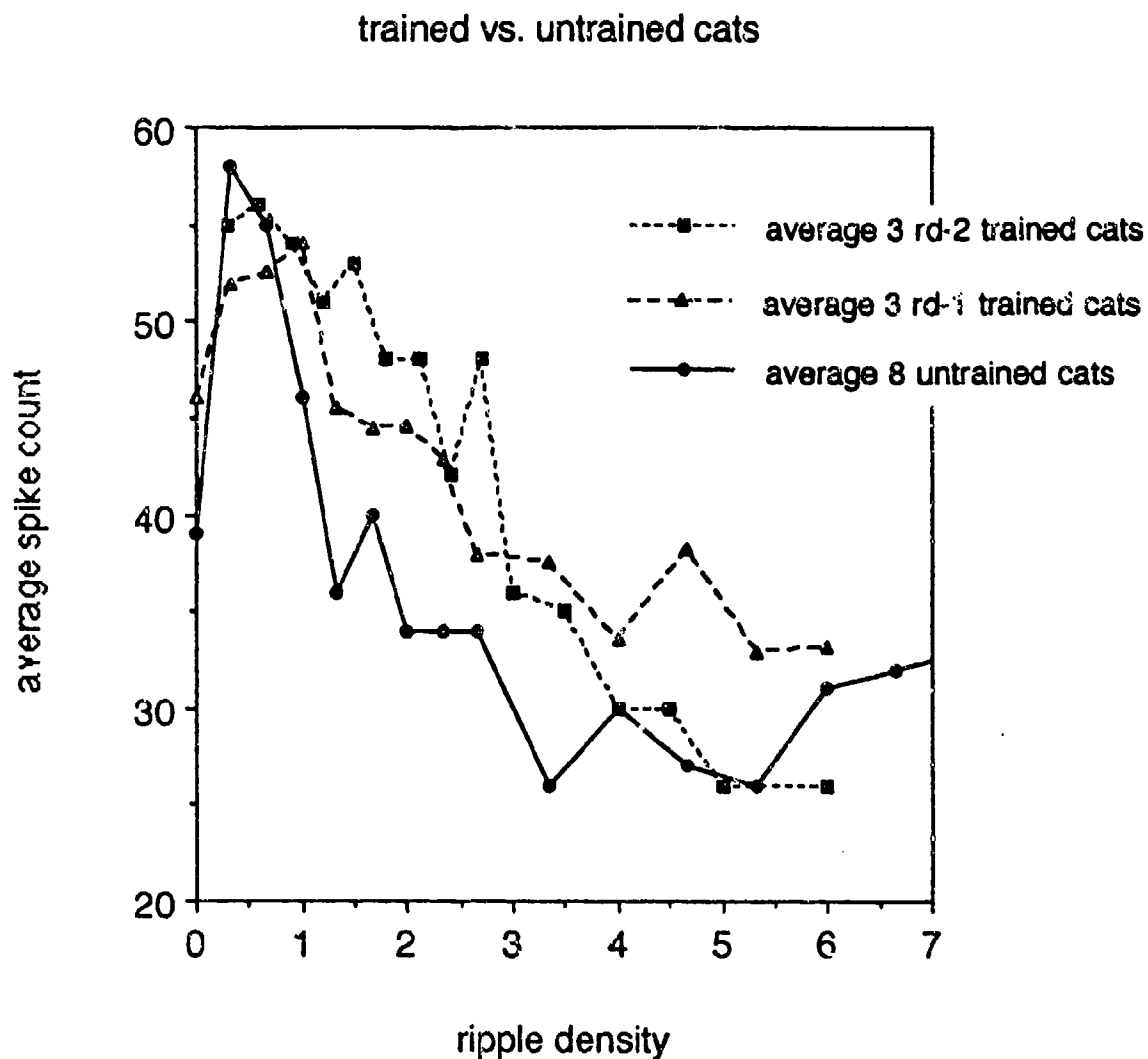
Figure 7. Comparison of Ripple Transfer Function for Trained Cats (rd = 1, AND rd = 2) vs. Untrained Cats

Tuning

The sharpness of tuning of the excitatory response areas of the penetrations for untrained and trained cats were measured and compared. The Q-40 value is calculated by dividing the characteristic frequency (cf) at the particular penetration by the bandwidth 40 dB above the threshold. Q-40 values for only penetrations with cfs below 11 kHz were included in this analysis.

Figure 8 displays the regression of the Q-40 values against the cf at each penetration for the untrained cats. Figure 9 displays the equivalent regression for the trained cats. In these figures, there are 139 values for the three trained cats, and 155 values for seven un-trained cats. In both cases, there is no relationship between the Q-40 value and the cf of the penetration, indicating that the sharpness of tuning is independent of the frequency of the penetration. .

$$y = -.016x + 2.054, \ r^2 = .001$$
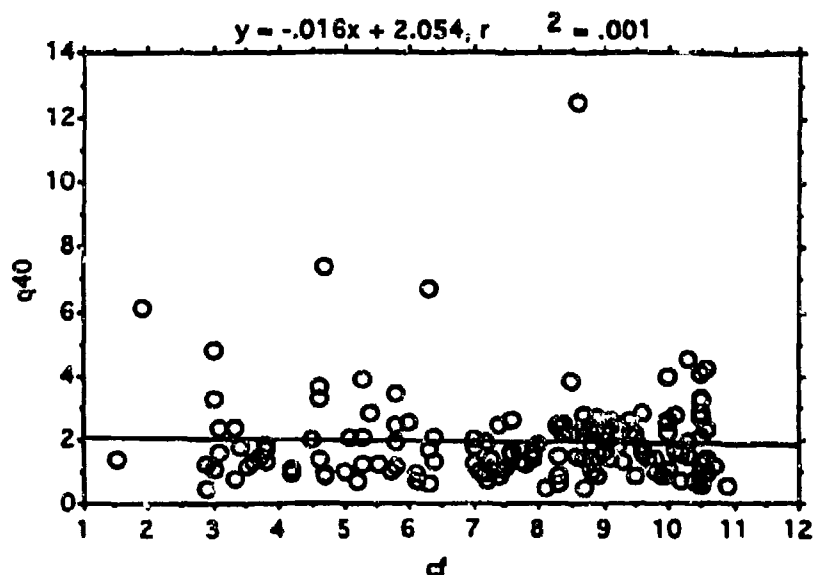
Figure 8. Regression of the Q-40 values against the cf at each penetration for the untrained cats.

$$y = -.034x + 2.083, \ r^2 = .001$$
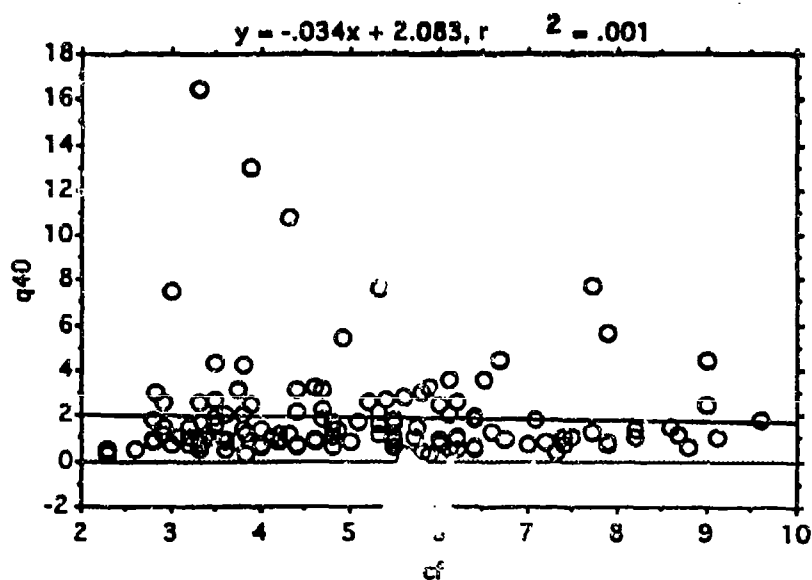
Figure 9. Regression of the Q-40 values against the cf at each penetration for the trained cats.

Figure 10 illustrates the distribution of the same Q-40 values for the three trained cats and for seven untrained cats, including 139 Q-40 values for the three trained cats, and 155 Q-40 values for seven un-trained cats. A Mann-Whitney U two-group un-paired comparison was performed; the two groups of animals, trained and untrained had different Q-40 values at a significance level of $p=0.004$. The mean Q-40 value for the untrained cats was 1.93, with standard deviation of 1.4; for the trained cats the mean Q-40 value was 1.91, with standard deviation of 2.2.
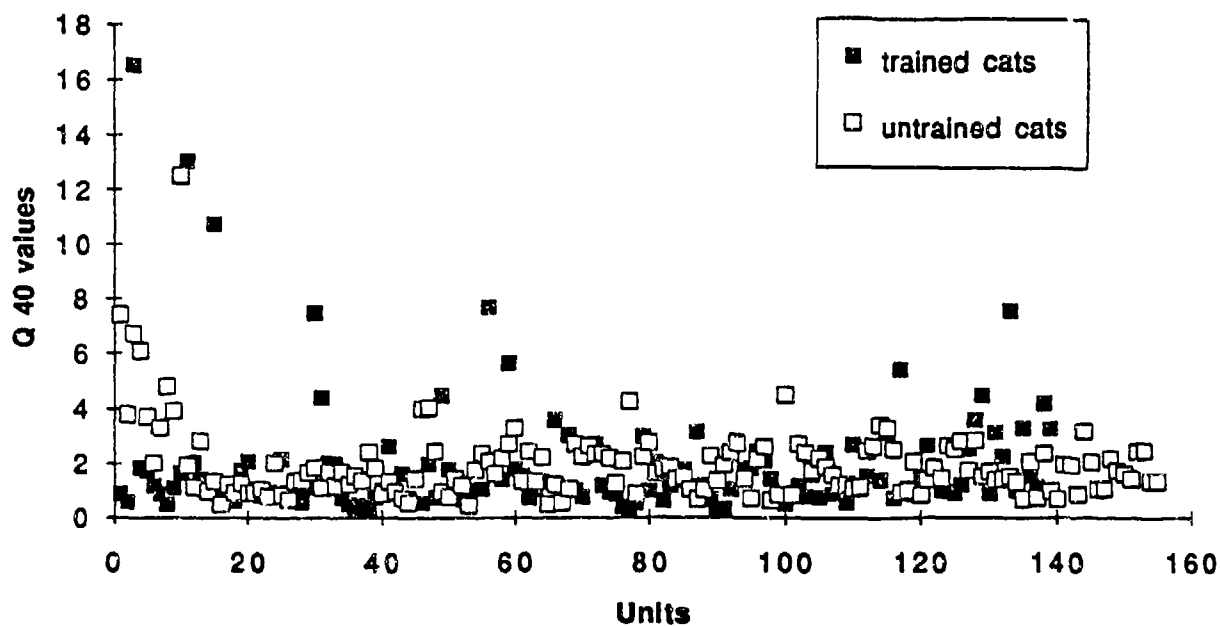
Figure 10. Distribution of Q-40 values for trained and untrained cats.
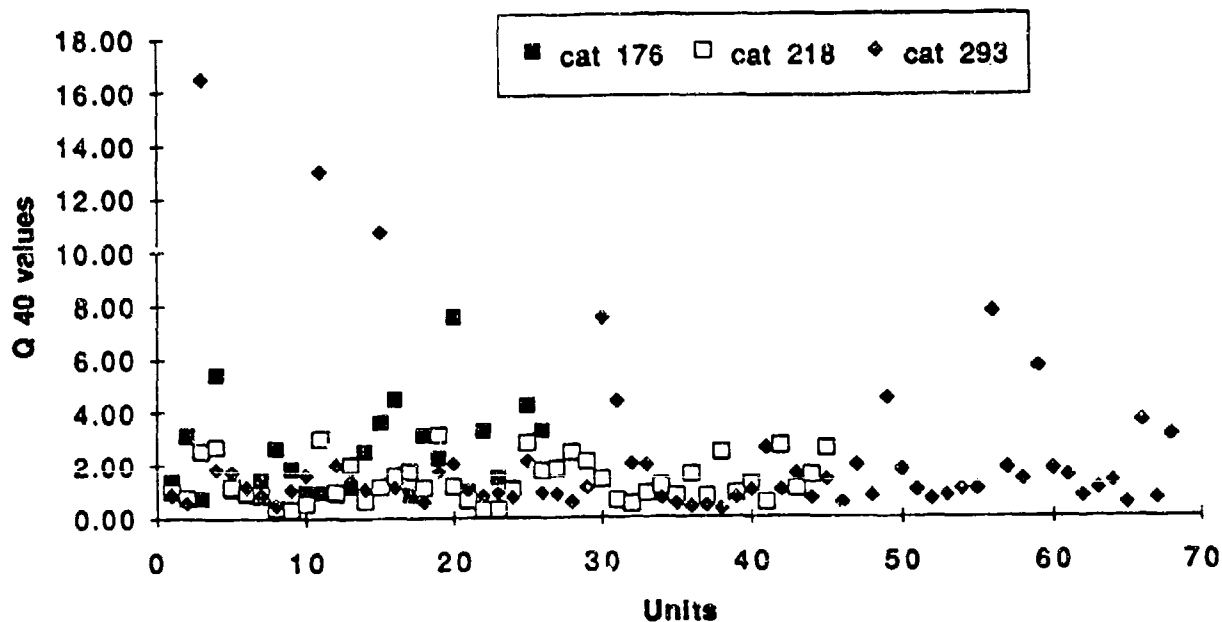
Figure 11. Distribution of Q-40 values for three trained cats.

Figure 11 illustrates the distribution of Q-40 values for the three trained cats. The cat with the best (lowest) ripple envelope phase shift threshold (18 degrees), #293, had the highest Q-40 values; the cat with a threshold of 33 degrees, #176, had the next highest Q-40 values; and the cat with the poorest threshold (80 degrees) had the lowest Q-40 values. The average Q-40 values and the behavioral thresholds for each of the trained cats are shown in the table below.

Cat	Q-40	behavioral threshold (in degrees)
218	1.37	80
293	2.09	18
176	2.36	33

Table 2. Q-40 values and behavioral thresholds for trained cats
The average Q-40 values for the two "exposed" cats, #122 and 40764, i.e. cats which had been exposed to the ripple stimulus for approximately two and four months respectively, but which were unable to learn the single speaker procedure, were 1.75 and 1.61. The correlation between the average Q-40 value for each of the three trained cats and its behavioral threshold was -.88, with an R-squared value of .77. A regression analysis of the trained cats' Q-40 values and their behavioral thresholds is shown in Figure 12. Due to the small sample size, the data do not allow a statistically secure p-value conclusion; however the R-squared value of 0.77 indicates that 77% of the variance in the data is explained by the relationship between the sharpness of tuning and the behavioral phase shift thresholds.
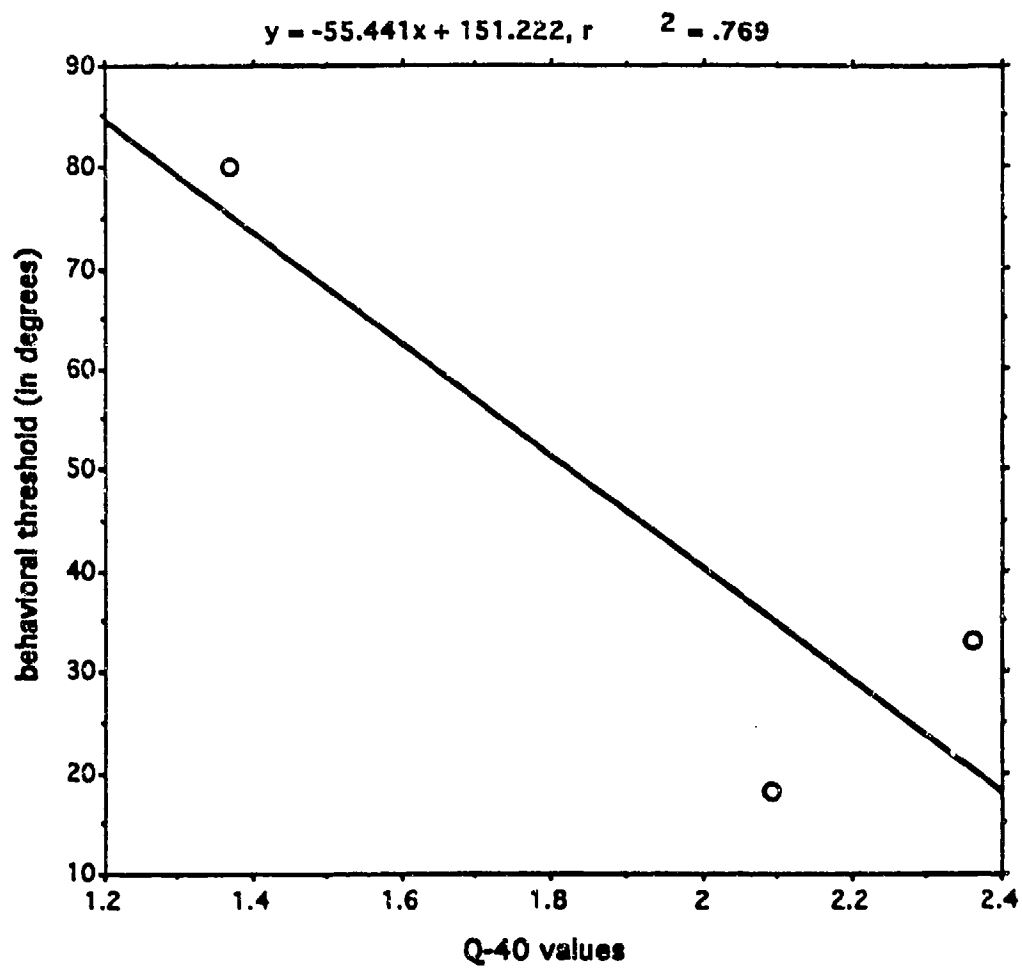
Figure 12. Regression analysis for the trained cats' Q-40 values and their behavioral thresholds.

Schreiner & Mendelson (1990) observed a frequency-independent maximum in sharpness of tuning for both Q-10 and Q-40 values located near the center of the dorsoventral extent of AI. A representative pseudo-three-dimensional projection of the spatial distribution of Q-40 values in AI, taken from their paper, is shown in Figure 13. In order to be certain that the Q-40 values of the trained cats were not penetrations recorded only at the very central portion of AI, three-dimensional representations of the spatial distribution of Q-40 values for the three trained cats were created. They are shown in Figure 14. The highest Q-40 values do not appear to be concentrated in the central portion of AI. This observation lends support to the idea that the training of the animals affected the sharpness of tuning, as measured by the Q-40 values.
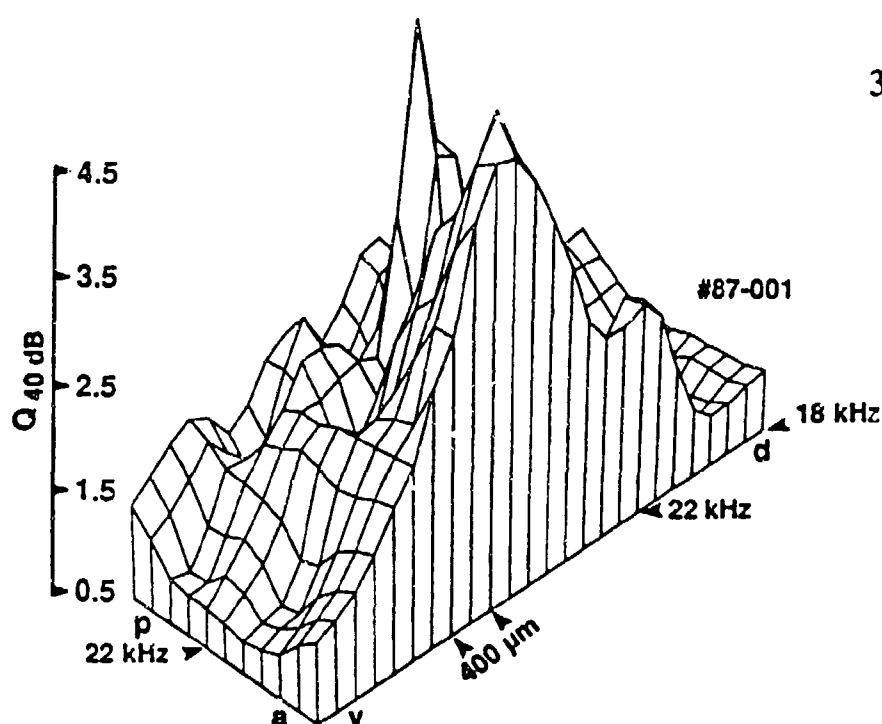
Figure 13. A representative pseudo-three-dimensional projection of the spatial distribution of Q-40 values in AI (from Schreiner & Mendelson, 1990). A maximum is seen in the center of the dorso-ventral extent.
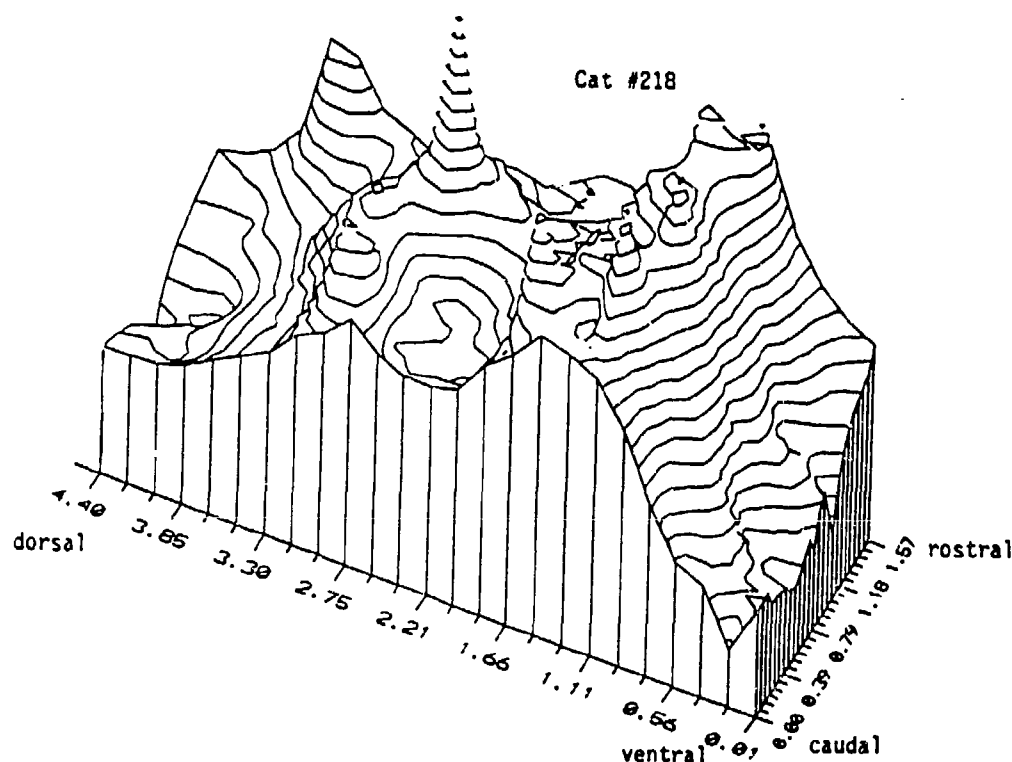
Figure 14 a). Three-dimensional representation of the spatial distribution of Q-40 values for trained cat #218.
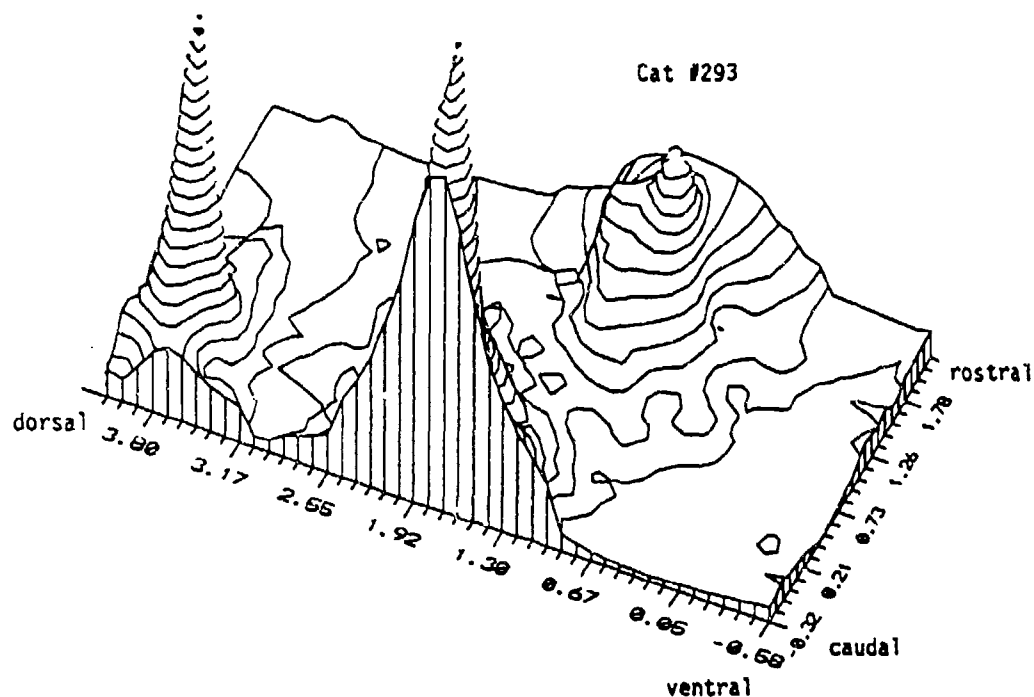
Figure 14 b) Three-dimensional representation of the spatial distribution of Q-40 values for trained cat #293.
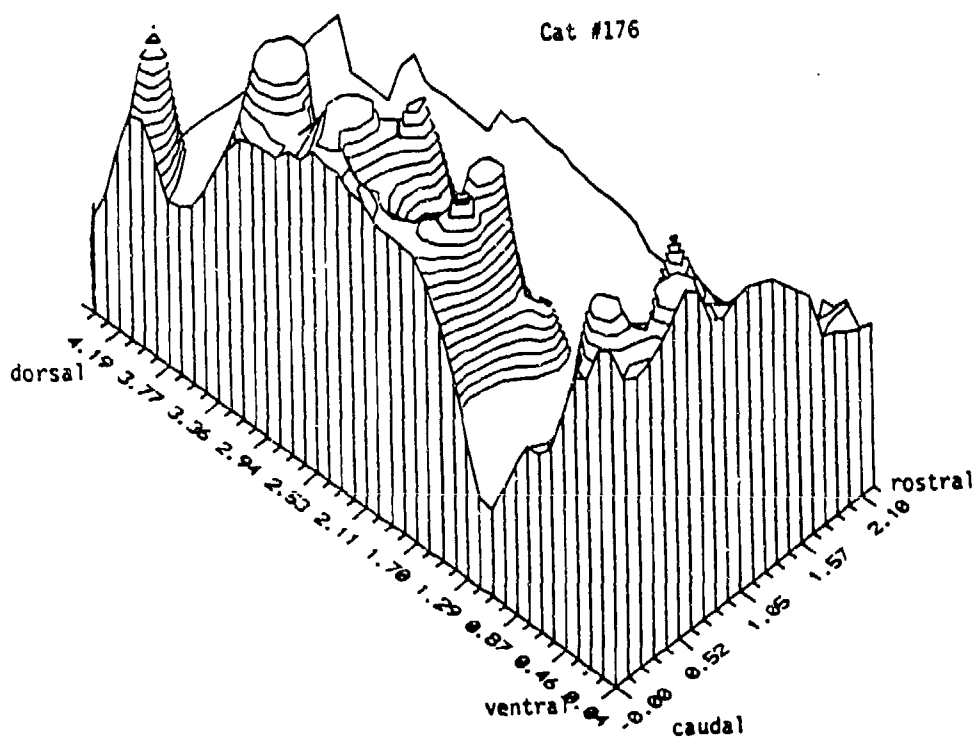
Figure 14 c) Three-dimensional representation of the spatial distribution of Q-40 values for trained cat #176.

DISCUSSION & CONCLUSIONS

The results of this research have several implications. The breadth of the subject matter bespeaks the need for an integration of several lines of investigation, to most effectively answer questions about brain function and perception and language.

Cortical Representation

Ripple densities were represented differently in the auditory cortices of the trained cats as compared to untrained cats. The peak of the ripple transfer function was shifted slightly toward higher ripple densities (1 ripple/octave compared to 0.33). The best ripple density peak response for untrained animals was somewhat lower than the mean best ripple density reported in Schreiner et al. (1993): 0.33 as compared to 1 to 2 ripples/octave. That may be partly due to the fact that most of the Schreiner et al. (1993) data were collected as single units, whereas the majority of the data reported for the untrained cats in the present study was collected as the response of small clusters of units (2-6 neurons), which may be less precise.

In hindsight, it would have been more telling to have trained the animals on a ripple density other than 1 ripple/octave, as untrained cats already show their best response to ripple densities in the range of about .33 to approximately 1 ripple/octave. However, it appears to be more difficult for cats to learn the discrimination at higher ripple densities; it was not possible to obtain a behavioral threshold at higher ripple densities.

The bandwidth 50% down from the peak was wider in trained cats than in untrained cats . A broadening of the filter function implies that a larger number of spatial frequencies (ripple densities) were processed

more efficiently by the trained cats as compared to the untrained cats. Since three trained cats were only trained on a ripple density of 1 ripple/octave, this increased processing efficiency was not due to a direct learning effect. What may have happened was an improvement in peak spacing discrimination that generalized somewhat so that spacings close to 1 ripple/octave, though not challenged, were affected in a facilitory manner.

Q-40 values over all penetrations were higher in trained cats as compared to untrained cats, indicating a general sharpening of frequency response tuning. The behavioral threshold was correlated (-.87) with Q-40 values, indicating that the sharpening in tuning was related to the proficiency of auditory discrimination. Recanzone et al. (1993) reported a sharpening of tuning for behaviorally relevant frequencies in monkeys trained to discriminate small differences in the frequencies of sequentially presented tonal stimuli. In this study, there were not enough penetrations with cf at the center frequency of the behavioral stimulus to compare the tuning there to the tuning at other frequency locations, nor would we expect a simple frequency contrast to code this discrimination. The fact that the overall tuning was sharpened may indicate a more global effect for this complex stimulus.

In the Recanzone study, the sharpness of tuning was not correlated with behavioral performance; however the cortical representation of the behaviorally relevant frequencies was increased and its area was correlated with behavioral performance. In the current study, a best ripple density response at each penetration was measured, and an increase in representation of best ripple density of 1 ripple/octave was not consistently seen in all trained animals as compared to untrained animals. However, although there is some indication of a spatial organization for ripple

density representation orthogonal to the frequency axis (Schreiner et al., 1993), this representation is not presently as circumscribed as is the representation for pure tones. A larger n of trained animals might reveal a reliable change in this representation.

Language learning

Kuhl, Williams, Lacerda, Stevens, & Lindblom (1992) observed that exposure to a specific language in the first six months of life alters infants' phonetic perception. Infants in America and Sweden were tested with both native- and foreign-language vowel sounds. "Infants from both countries exhibited a language-specific pattern of phonetic perception" (p.606). As pointed out, this perceptual organizing occurs well before the age at which children begin to acquire word meanings, i.e. well before the acquisition of language. It was suggested that "linguistic experience shrinks the perceptual distance around a native-language prototype, in relation to a nonprototype, causing the prototype to perceptually assimilate similar sounds" (p.608). Thus phonetic prototypes were conjectured to be "fundamental perceptual-cognitive building blocks" (p.608).

The data in this report support the idea that certain basic auditory processing capabilities are operative in humans and in other mammals as well, specifically at least in cats, and that exposure to or experience with the stimuli refines the processing capability and the cortical representation of that stimulus. In the same way that a perceptual organization was set up in the brains of the six-month-old infants, so categorization and learning occur during the acquisition of a first or second language. Learning occurred in these animals as they performed the auditory discrimination, and their brain representations of these behaviorally important stimuli

were changed in parallel. It seems plausible that, at birth, certain basic auditory perceptual capabilities are present, and that with linguistic exposure, certain contrasts are sharpened and certain other unused phonetic contrasts drop out of the discriminative repertoire. The cortical organization thus dynamically represents those stimuli to which the human/animal has been exposed and which are informationally and biologically relevant to that creature.

Evolution of linguistic abilities

Clearly humans are more "intelligent" than cats. If in doubt, try training a cat on an operant discrimination procedure. Cats were able to learn the paradigm, however, and showed quite fine resolution abilities (at least two out of three cats did), but there was variability in performance, and two other cats were unable to learn the discrimination procedure. Thresholds for the two adept cats were about 3 - 4 times higher than humans for the equivalent ripple density. Their intelligence as well as their perceptual discrimination abilities for detecting envelope phase shifts are at a lower level but are on the same continuum as humans.

Watson (1991) has recently reported significant relationships between simple sensory, cognitive, and motor abilities and psychometric intelligence. She reminds us that Spearman (1904) and others reported relationships between pitch discrimination and intelligence. The suggestion is that working memory is a limited-capacity system, and that therefore individuals who can process information rapidly will have more working memory available to process new information. Thus, the sensory ability to perceive differences and make discriminations can be considered part of a processing capacity. Learning calls upon this information-processing ability.

The innate abilities with which humans are born interact with experiential input, especially during development but later as well, to produce a cerebral organization which reflects that interaction. The huge amount of neural modification that occurs within the brain during development allows us to "make sense of the world". As Edelman (1987) says: "it is difficult to imagine the world as it is presented to a newborn organism...the environment presented [to such an organism] is inherently ambiguous; even to animals eventually capable of speech such as ourselves, the world is initially an unlabeled place" (p.3). Adaptive behavior requires initial categorization of salient aspects of the environment so that learning can occur (p.4). Thus, perceptual abilities are very important in order to form concepts. All animals are born with certain basic sensory and discriminative capacities; all must interact successfully with their environment in order to survive. This capacity for learning and for change has allowed Homo sapiens to evolve. In turn, Homo sapiens has developed speech and language, culture and community. Our capacity for speech has advanced our development of thought, which has changed the world. These faculties have evolved over many eons of time due to the mechanism of neural adaptation. How this cortical change occurs is unknown. Evidence that it occurs is amazing in itself.

ACKNOWLEDGMENTS

REFERENCES

Allard, T., Clark, S.A., Jenkins, W.M., & Merzenich, M.M.
Reorganization of somatosensory area 3b representations in adult owl
monkeys after digital syndactyly. Journal of neurophysiology, 1992, Sep.,
66.3, 1048-1058.

Chistovich, L.A. Central auditory processing of peripheral vowel spectra.
Journal of the Acoustical Society of America, 1985, March, 77(3), 789-
805.

Clark, S.A., Allard, T., Jenkins, W.M., & Merzenich, M.M. Receptive
fields in the body-surface map in adult cortex defined by temporally
correlated inputs. *Nature*, 1988, 332, 31 March, 444-5.

Edelman, G.M. *Neural Darwinism - the theory of neuronal group
selection*. New York: Basic Books, 1987.

Harrison, R.V., Nagasawa, A., Smith, D.W., Stanton, S., & Mount, R.J.
Reorganization of auditory cortex after neonatal high frequency cochlear
hearing loss. *Hearing Research*, 1991, 54, 11-19.

Heil, P. & Scheich, H. Functional organization of the avian auditory cortex
analogue. I. Topographic representation of isointensity bandwidth. *Brain
Research*, 1991, Jan.18, 539, 1, 110-120.

Imig, T.J., & Adrian, H.O. Binaural columns in the primary field (AI) of
cat auditory cortex. *Brain research*, 1978, 138, 241-57.

Imig, T.J., & Brugge, J.F. Sources and terminations of callosal axons
related to binaural and frequency maps in primary cortex of the cat.
Journal of comparative neurology, 1978, 182, 637-60.

Jenkins, W.M. & Merzenich, M.M. Reorganization of neocortical
representations after brain injury: A neurophysiological model of the bases
of recovery from stroke. *Prog. Brain Research*, 1987, 71, 249.

Jenkins, W.M. & Merzenich, M.M. Role of cat primary auditory cortex
for sound-localization behavior. *Journal of neurophysiology*, 1984, 52, 5,
819-47.

Jenkins, W.M., Merzenich, M.M., Ochs, M.T., Allard, T., & Guic-Robles, E. Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *Journal of neurophysiology*, 1990, Jan, 63, 1, 82-104.

Keeling, M.D., Jenkins, W.M., & Schreiner, C.E. Discrimination of vowel-like stimuli in the cat. 1994.

Keeling, D., Schreiner, C.E., & Jenkins, W.M. Discrimination of formant shifts in vowel-like ripple spectra. *Association for Research in Otolaryngology Abstracts*, 1992, 201.

Knight, P.L. Representation of the cochlea within the anterior auditory field (AAF) of the cat. *Brain research*, 1977, 130, 447-67.

Kuhl, P.K. Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 1992, Jan.31, 255, 606-608.

Langner, G. Bonke, D., & Scheich, H. Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. *Exp. Brain Research*, 1981, 43, 11-24.

Lass, N.J. *Handbook of speech-language, pathology, and audiology*. Toronto: B.C. Decker, 1988.

Lloyd, R.J. *Some Researches into the Nature of Vowel-Sound*. Liverpool, England: Turner and Dunnett, 1890.

Mendelson, J.R., Schreiner, C.E., Grasse, K., & Sutter, M. Spatial distribution of responses to FM sweeps in cat primary auditory cortex. *Association for Research in Otolaryngology Abstract*, 1988, 11, 36.

Mendelson, J.R., & Grasse, K.L. A comparison of monaural and binaural responses to frequency modulated (FM) sweeps in cat primary auditor cortex. *Experimental Brain Research*, 1992, 92, 1, 435-454.

Merzenich, M.M., Andersen, R.A., & Middlebrooks, J.H. Functional and topographic organization of the auditory cortex. In *Hearing mechanisms and speech, Experimental brain research, suppl. II*, O. Creutzfeld, H. Scheich, & Chr. Schreiner (Eds.). Berlin: Springer-Verlag, 1979.

Merzenich, M.M., Knight, P.L., & Roth, G.L. Representation of cochlea within primary auditory cortex in the cat. *Journal of neurophysiology*, 1975, 38, 231-49.

Merzenich, M.M., Roth, R.A., Knight, P.L., & Colwell, S.A. Some basic features of organization of the central auditory system. In *Psychophysics and physiology of hearing*, E.F. Evans & J.P. Wilson (Eds.). London; Academic Press, 1977.

Middlebrooks, J.C., Dykes, R.W., & Merzenich, M.M. Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. *Brain research*, 1980, 181, 31-48.

Miller, J.D. Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America*, 1989, May, 85(5), 2114-2133.

Newman, J.D., & Wollberg, Z. Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain research*, 1973, 54, 287-304.

Peterson, G.E., & Barney, H.L. Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 1952, 24, 175.

Recanzone, G.H., Jenkins, W.M., Hradek, G. T., & Merzenich, M.M. Progressive improvement in discriminative abilities in adult owl monkeys performing a tactile frequency discrimination task. *Journal of Neurophysiology*, 1992, May, 67, 5, 1015-30.

Recanzone, G.H., Merzenich, M.M., Jenkins, W.M., Grajski, K.A., & Dinse, H.R. Topographic reorganization of the hand representation in cortical area 3b of owl monkeys trained in a frequency discrimination task. *Journal of Neurophysiology*, 1992, 67, 1031-56.

Recanzone, G.H., Schreiner, C.E., & Merzenich, M.M. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience*, 1993, Jan. 13, 1, 87-103.

Reale, R.A., & Imig, T.J. An orderly frequency representation in the posterior ectosylvian sulcus of the cat. *Neuroscience Abstr*, 1977, 3, 10.

Robertson, D. & Irvine, D.R.F. Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *Journal of Comparative Neurology*, 1989, 282, 456-71.

Rose, J.E. The cellular structure of the auditory region of the cat. *Journal of comparative neurology*, 1949, 91, 409-39.

Schreiner, C.E., Calhoun, B., & Keeling, D. Physiology and topography of cortical neurons explored with vowel-like ripple-spectra. *Association for Research in Otolaryngology Abstract*, 1993, 161.

Schreiner, C.E., & Cynader, M.S. Basic fundamental organization of second auditory cortical field (AII) of the cat. *Journal of Neurophysiology*, 1984, 51, 6, 1284-1305.

Schreiner, C.E., & Mendelson, J.R. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *Journal of Neurophysiology*, 1990, Nov., 64, 5, 1442-1459.

Schreiner, C.E., Mendelson, J.R., & Sutter, M.L. Functional topography of cat primary auditory cortex: representation of tone intensity. *Experimental Brain Research*, 1992, 92, 1, 105-122.

Schreiner, C.E., & Urbas, J.V. Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). *Hearing research*, 1986, 21, 227-41.

Schreiner, C.E., & Urbas, J.V. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hearing research*, 1988, 32, 49-64.

Shamma, S.A., Fleshman, J.W., Wiser,P.W., & Versnel, H. Organization of response areas in ferret primary auditory cortex. *Journal of Neurophysiology*, 1993, Feb. 69,2, 367-383.

Spearman, C. "General intelligence," objectively determined and measured. *American Journal of Psychology*, 1904, 15, 201-293.

Steinschneider, M., Arezzo, J.C., & Vaughan, H.G.Jr. Tonotopic features of speech-evoked activity in primate auditory cortex. *Brain Research*, 1990, 519, 158-168.

Suga, N. Auditory neuroethology and speech processing: complex-sound processing by combination-sensitive neurons. In *Auditory function - neurobiological bases of hearing*, G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.). New York: John Wiley & Sons, 1988.

Suga, N. The extent to which biosonar information is represented in the bat auditory cortex. In *Dynamic aspects of neocortical function*, G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.). New York: John Wiley & Sons, 1984.

Watson, B.U. Some relationships between intelligence and auditory discrimination. *Journal of Speech and Hearing Research*, 1991, June, 34, 621-627.

Winer, J.A., Diamond, I.T., & Raczkowski, D. Subdivisions of the auditory cortex of the cat: the retrograde transport of horseradish peroxidase to the medial geniculate body and posterior thalamic nuclei. *Journal of comparative neurology*, 1977, 176, 387-417.

Wollberg,A., & Newman, J.D. Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science*, 1972, 175, 14 Jan., 212-4.

Woolsey, C.N., & Walzl, E.M. Topical projection of nerve fibers from local regions of the cochlea to the cerebral cortex of the cat. *Bulletin of Johns Hopkins Hospital*, 1942, 71, 315-44.

DISCRIMINATION OF VOWEL-LIKE STIMULI IN THE CAT

M. Diane Keeling [a], William M. Jenkins, and Christoph E. Schreiner, Coleman Memorial Laboratory, W.M. Keck Center for Integrative Neuroscience, University of California, San Francisco; San Francisco, CA 94143-0732.

[a] now at Psychology Dept., Box 5050, University of New Brunswick, Saint John, Saint John, N.B. E2K 2E2, Canada.

Abstract

The objective of this study was to investigate the perception of vowel-like stimuli by measuring the ability of cats to detect changes in frequency locations. Three cats were trained and tested for their ability to discriminate frequency peak shifts of a broad-band spectrally-modulated harmonic complex, centered at 3.675 kHz, with a fundamental frequency of 69 Hz, in a 2-Alternative Forced Choice procedure. The time required to train to the procedure varied from 3 weeks to 2 months. The threshold (75% correct) was lower for all three cats at the end of testing, as compared to the threshold at the beginning of training. Performance improvements were continuous and progressive. The behavioral thresholds correlated with the electrophysiologically-measured Q-40 values of overall neural responses (described in a subsequent paper), indicating that a sharpening in tuning was related to the degree of proficiency in the auditory discrimination. The data imply that language acquisition involves dynamic cortical reorganization of speech feature representations and that human linguistic capabilities evolved from basic mammalian auditory processing capacities.

It is generally believed that the vocal and auditory systems have co-evolved so that "the vocal system has adapted to produce sounds suitable for detection and processing by the auditory system, and the auditory system has evolved to detect and process these sounds" (Suga, 1988, p.684). The development of language in humans may have taken advantage of acoustic parameters that accentuate certain aural differences just as vowels are discriminated using formants - those frequencies which are boosted by the particular configuration of the vocal tract. It may also be the case that certain acoustic cues common to speech and to other biologically significant communicative sounds are processed by basic neural mechanisms that are similar in humans and in other animals.

Partly because of the complexity of speech perception, Liberman, Cooper, Shankweiler, & Studdert-Kennedy, (1967) and others (e.g. Chomsky, 1972) have suggested that "speech is special", i.e. that humans have unique linguistic capacities designed specifically for producing, perceiving, and understanding speech and language. Furthermore, it has been suggested that the perception and production of speech are intimately related in that the perception of speech is directly dependent upon neural "knowledge" of its production. In this respect, speech perception is considered to be qualitatively different from other forms of auditory communication.

Animal Communication

There has been considerable discussion as to whether communication among animals is comparable to communication among humans. The discussion has revolved around defining what is meant by "communication" vs. "language" (see for example Tartter, 1986, chapter 8; and Lieberman, 1984). Attempts have been made to train animals (mainly apes, gorillas,

and chimpanzees) to use human language. These studies were unable however, to differentiate between cognitive abilities, that is, abilities to form associations between concepts/objects and an ability to form a truly linguistic symbolic representation of an object/idea. (See the following references: Gardner & Gardner, 1969; Patterson, 1978 a, b; Premack, 1971; Rumbaugh, Gill, & von Glaserfield, 1973; and Terrace, Pettito, Sanders, & Bever, 1979.)

Field data for vervet monkeys in their natural environment in Kenya were presented by Seyfarth, Cheney, & Marler in 1980. The animals gave different alarm calls to different classes of predators and reacted to play-backs in different ways, indicating that the calls carried distinctive semantic meanings. Cheney & Seyfarth (1988) further showed that monkeys who had learned to ignore an unreliable signaler (a monkey "crying wolf"?) also ignored an acoustically different but same referent call from the same individual. Such transfer did not occur, however, if the identity of the signaler or the referent of the call changed. The authors suggest that vervet monkeys, "like humans, process information at a semantic, and not just an acoustic level" (p.477).

Discrimination of specific speech features was reported by Kuhl and Miller in 1975, in the chinchilla (a mammal with auditory capabilities fairly similar to those of humans - see Miller, 1970). Using a shock avoidance procedure, they trained animals to respond differently to a variety of spoken /t/ and /d/ consonant-vowel syllables. Once trained, the animals generalized to other speakers, other vowel contexts, and to computer-synthesized /ta/ and /da/ syllables.

In another series of investigations, Kuhl & Padden (1982, 1983) tested monkeys, using a positive reinforcement technique, wherein the

monkey was trained on a same-different task, i.e. the monkey was rewarded for one behavior when stimuli were the same, and for another when the stimuli were different. Voice-onset-time (VOT), defined as the time between the stop release burst and the onset of vocal cord vibration (voicing), and place-of-articulation discrimination were investigated. In both cases, discriminability was poor when the stimulus pairs were from the same phonetic category and good when they were from different phonetic categories.

Other investigators have reported discrimination of speech parameters in animals. For example, Morse & Snowdon (1975) reported categorical discrimination of speech sounds by rhesus monkeys using the cardiac component of the orienting response. The monkeys significantly differentiated stimuli across phonetic boundaries, but also discriminated within phonetic categories, although to a lesser degree. Sinnott, Beecher, Moody, & Stebbins (1976) concluded that monkeys and humans had similar sensory capacities; however, while humans showed a longer latency of response when both stimuli were within the same phonemic category, monkeys did not. They argued that this was support for unique speech processing capacities in humans. Sinnott & Adams (1987) measured difference limens to various VOT stimuli in monkeys and humans and observed that again, monkeys were much less sensitive to speech cues than were humans. Waters & Wilson (1976) reported perceptual boundaries in rhesus monkeys that were close to the human boundaries between voiced and voiceless consonants.

Baru (1975) discussed the parameters involved in discrimination of vowels in the dog. Dogs were able to discriminate /a/ from /i/ regardless of fundamental frequency (male voice vs. female voice), and discrimination

of vowels appeared to rely on the same cues as those for humans (mainly the first two formant frequencies of the signal). Following either unilateral or bilateral ablation of primary auditory cortex, discrimination was not impaired except when signal duration was 30 msec (as opposed to 75, 150, or 300 msec).

Dewson (1964) reported vowel discrimination in the cat. Cats were trained to discriminate between /u/ and /i/, but following bilateral ablation of the ventral insular-temporal cortex (but not all of primary auditory cortex) did not retain or were unable to relearn the discrimination. The animals could still discriminate frequency and intensity differences.

Vowel formants

Vowels are characterized by their formants - frequencies which are elevated in amplitude relative to other frequencies and which are produced by the natural resonances of the vocal tract activated during production of the vowel (indicated as F1, F2, F3, etc.). Formant values show considerable variability across speakers, however, and even overlap for some vowels (Peterson & Barney, 1952). The variability across speakers is due to physical differences in vocal tract size but also to a speaker's past experience, mainly his/her particular dialectical background, but the values may vary within context and even across repetitions as well.

It is clear that in spite of speaker variation, listeners are somehow able to adjust for this. Miller (1989) reports that in the late 1800's Richard John Lloyd formulated the formant-ratio theory of the vowel stating that "vowel quality depends on the intervals between the resonances, not their absolute values" (p.2155). Miller's auditory-perceptual interpretation of the vowel is a descendant of the formant-ratio theory and proposes that

changes in spectral *patterns* of formants effectively categorize vowel sounds.

The ratios of the formants may be an important cue. We calculated the ratios of the first and second, and second and third formants; since frequency is represented on a logarithmic scale along the basilar membrane within the cochlea, we computed the ratios of the formants in units per octave (Miller,1989) and hypothesized that the perception of variations in the position of spectral peaks is dependent on the spacing of the peaks.

Cat vocalizations have been analyzed and appear to contain spectral peaks of energy similar to formants in human vowels (Brown, Buchwald, Johnson, & Mikolich, 1978). The average fundamental frequency for cats is within the range of 400-600 Hz (Shipley. Carterette, & Buchwald, 1990); it was also observed that the frequency with the largest energy is usually a resonance of the vocal tract. The authors state: "the presence of formants in cat vocalizations and the magnitude of energy in these frequencies suggests that in processing conspecific vocalizations the auditory system of the cat needs to perform extensive identification and tracking of formant patterns" (p.15).

Thus, it can be concluded that the perception of vowels involves the processing of peaks of spectral energy, the frequencies of which are located at various distances or spacings. In order to elucidate the precision of this processing, vowel-like stimuli were presented to cats in a 2-Alternative Forced Choice testing procedure. Holding the peak spacing/ripple density constant, in each test condition, the phase of the envelope of the stimulus was shifted so that the locations of the peaks changed. The precision with which the peak shifts could be determined was expected to shed light upon the importance of small changes in formant location for vowel processing.

METHODS

Three cats were trained to detect shifts in the phase of the stimulus envelope. A two-alternative forced choice paradigm was used. All procedures were approved by the University Committee on Animal Research (protocol # A886-05785-03). The stimulus generation and response tabulation were controlled and presented with the LabVIEW application.

The animals were food deprived to 90- 95% of their pre-fast weight; reward during testing was diluted canned cat food (Hill's Prescription Diet feline c/d). The food was dispensed from a large dose syringe using a stepping motor, which pumped a small amount of food onto a fiberglass plate, about 6 cm in diameter, which the cat licked. The food reward system was modelled after a dispensing system devised by Thompson, Porter, O'Bryan, Heffner & Heffner (1990). The cats were weighed daily and received a supplement of dry cat food after testing in order to obtain their total daily nutritional requirement. The test cage was located within a sound-attenuating chamber. Behavior was monitored continuously with a video camera.

Training procedure

Initially, a hungry cat was placed in the test cage, allowed to explore its new environment, and received food rewards under the investigator's control on a small food reward plate. In the next stage, the cat was trained to press a nose key located at mid-line to indicate an observing ("ready") response, for which it received food. The observing response not only ensured that the cat was attending and could initiate trials at its own pace and volition, but also forced the cat's head position to be constant from trial to trial. A center key press automatically triggered a food reward.

Alternatively, if the cat required "shaping", the investigator could produce a food reward as the cat progressively came closer and closer to pressing the center key on its own. For instance, if the cat oriented its head toward the center key, or moved toward the center key, or lightly pressed the center key, the investigator would deliver a food reward, requiring over trials behavior that came closer and closer to the desired autonomous behavior.

Once the cat was reliably pressing the center key in order to receive the food reward, auditory stimuli were introduced; reward then required a center key press, which was followed by auditory stimulus presentation, to which the cat was required to make a response either to the left or to the right side. Responses were indicated on two spatially separated nose-press keys, one to each side of the middle observing key. (See Figure 1.) To facilitate this stage of learning, the investigator usually remained present in the sound booth and baited the side keys with a small amount of food, by inserting a small narrow plastic tube with food on the end into the cage and onto the response key, which the cat readily followed. Eventually, by indicating the correct response with only a touch to or movement toward the correct key, the investigator could help the cat acquire this behavior. Shaping, by giving reward as the animal approached the correct response, on the control panel outside the booth, was also later used to assist the learning of this stage by the animal.

Initially, stimuli were presented from two small lateralized speakers, which were located at approximately 60 to 80 degrees to the left and right of midline. At the beginning of training the stimulus on the left side was either two or three stimuli, of the same phase (always 0), (designated AAA or AA); the stimulus on the right was maximally different, i.e. 180 degrees

different in phase from the left stimulus (designated BBB or BB). Over time the stimuli were changed to be AA on the left, and AB on the right. The requirement that the animal go left for reward to a left-side stimulus presentation and go right to a right-side stimulus presentation was usually fairly easily learned by the cats and took approximately 2 to 3 weeks of training.

As the animal learned to go left or go right based upon the stimulus presented, the speakers were moved in to the center location gradually until finally the ripple stimuli were presented from a single speaker overhead. Stimulus pair AA (i.e. two identical stimuli) required Response A (left-side key press) in order to receive reward; Stimulus pair AB (i.e. second stimulus different from the first) required Response B (right-side key press) in order to receive reward. The switch from discrimination based upon side of stimulus presentation to discrimination based upon the quality of the stimuli, that is, two stimuli the same vs. a second stimulus different from the first, was quite difficult for the cats to acquire. Some animals were unable to learn this task, and their training was discontinued. The three cats which were able to learn this discrimination varied in length of time required to learn the procedure from a minimum of 3 weeks to a maximum of two months. Performance, i.e. percent correct overall, and percent correct on the left and percent correct on the right, was calculated on-line, as for human subjects. Unlike the human testing however, there was no pre-set number of trials per session; cats continued working until they were sated and did not initiate any further trials. They usually performed anywhere from 100 to 250 trials per day. At the end of the test session, the stimulus phase and response values were stored. Thresholds were computed later, in the same manner as that used for the human

subjects, i.e. by a linear regression formula applied to the psychometric function which calculated the phase shift correctly detected on more than 75% of the trial presentations.

Stimulus

Stimuli were ripple spectra centered at 3.675 kHz, with a ripple density of 1.00 ripple/octave, a fundamental frequency of 69 Hz and a ripple bandwidtn of 3 octaves. The stimulus extended from 910 Hz to 7,153.2 Hz. The center frequency of 3.675 kHz was not so low as to preclude the possibility of recording, since frequencies lower than approximately 500 Hz to 1 kHz are usually located within the posterior ectosylvian sulcus and difficult to record from. Also, cat vocalizations appear to contain pea'cs of spectral energy, the largest being usually a resonance of the vocal tract and located between about 1 and 2 kHz, with spectral peaks extending up to approximately 4 kHz (Shipley, Carterette, & Buchwald, 1990). Thus, the center frequency of 3.675 kHz is a frequency likely to be of interest to cats, and possible to record from on the cortical surface (see also Heffner & Heffner- the hearing range of the cat, 1985). The phase of the envelope of the stimulus was shifted, originally by 180 degrees, in a positive direction only (a downward shift in the frequency of the center peak); the range of values was narrowed and lowered gradually as the animal's performance improved. The ripple density of the ripple stimuli was constant at one ripple/octave; thus, the distance between the ripple peaks remained the same, however the frequency locations of the peaks changed with the shift in ripple phase. Depth of modulation of the envelope was 30 dB. The intensity of the stimuli varied randomly by +/- 3 dB around approximately 64 dB SPL. Figure 2 illustrates the ripple

stimulus used for animal testing. At the end of testing, the cats underwent electrophysiological recording.
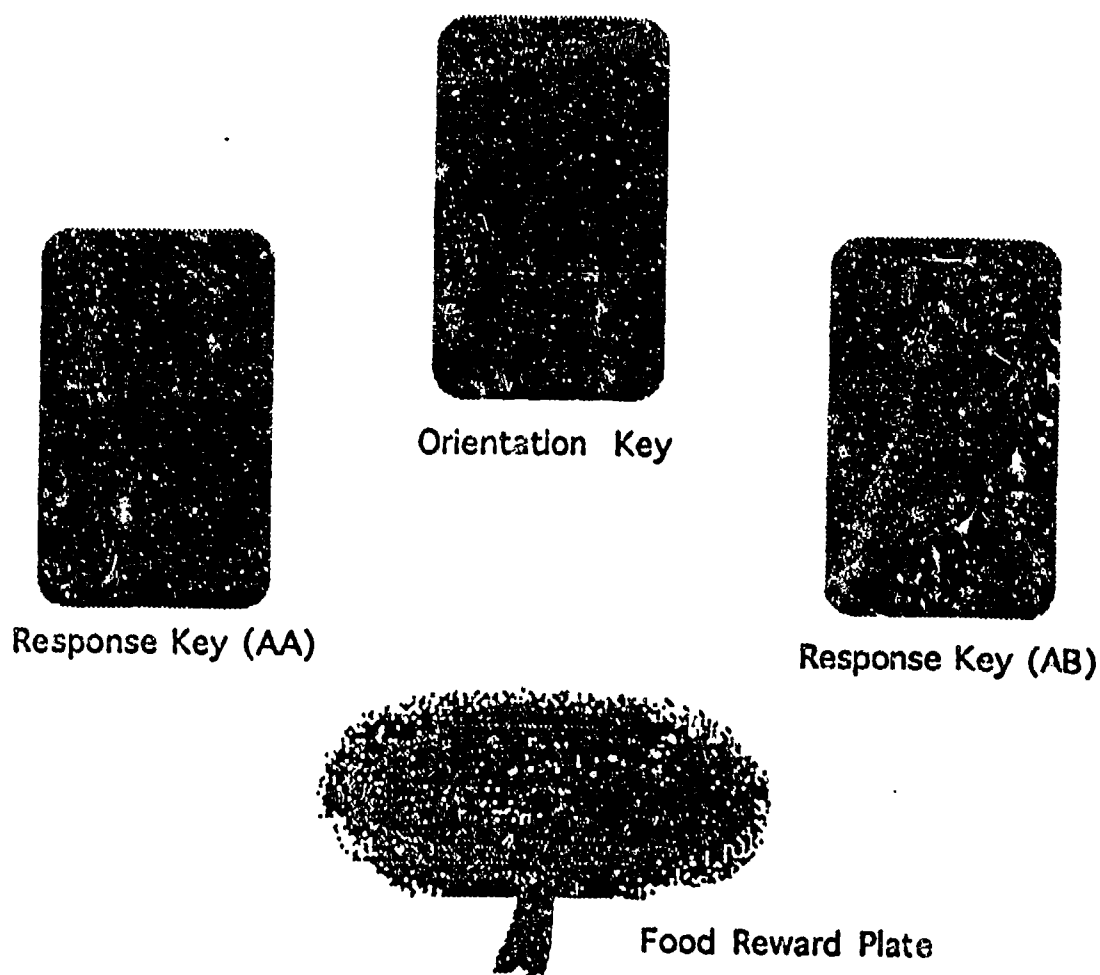
Figure 1. **Configuration of Animal Response Keys and Reward Plate. If the two stimuli were the same, the cat pressed the left response key to receive reward; if the second stimulus was different, the cat pressed the right response key to receive reward.**
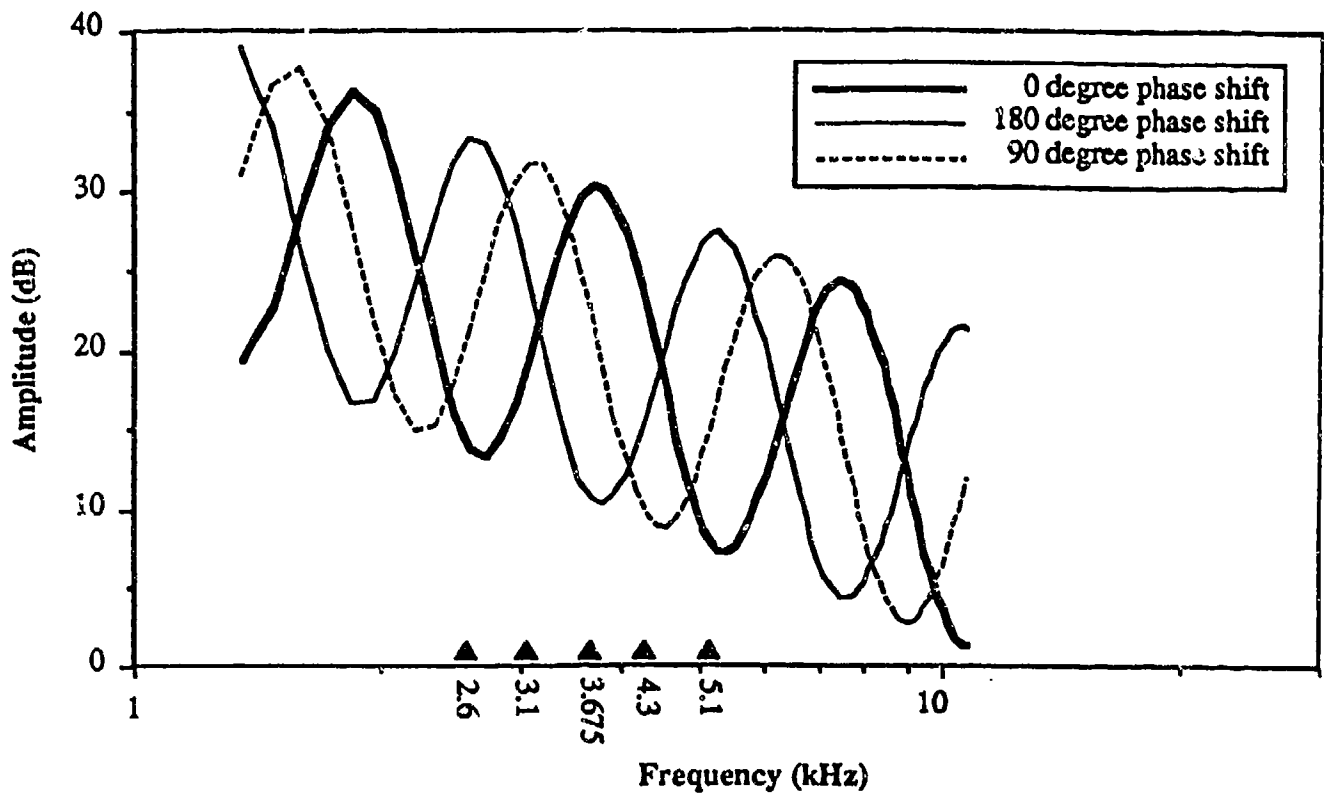
Figure 2. Ripple Stimulus for Animal Testing, showing 0
degree, 180 degree, and 90 degree phase shift. Corresponding
frequency locations are shown: 3.675, 2.6, and 3.1 kHz.
Frequencies to the right represent phase shifts of negative 90
and negative 180 degrees, at locations of 4.3 and 5.1 kHz.

RESULTS

The first behavioral cat (#218) began training at age ten months. It required almost two months to shape the cat to reliably perform detection of ripple envelope phase shifts. The cat was subsequently tested over a period of five months before it underwent electrophysiological recording. Threshold was computed after each daily session as the smallest degree positive shift in phase of the envelope of the stimulus that the cat could detect with 75% or greater accuracy. An example of a psychophysical function for this cat is shown in Figure 3. Its threshold in this particular test session was approximately 57 degrees.

Figure 4a shows the performance of the animal over the time course of the test sessions. The average threshold over the first ten test sessions was about 127 degrees. The final threshold for this animal (its average over the last ten test sessions) was a phase shift of 80 degrees.

The second trained cat (#293) began training at age eight months; it required only 3 weeks to train to the procedure. It was tested over a six-month period, and then underwent electrophysiological recording. Figure 4b shows the improvement in threshold over time for this cat. The average threshold over the first ten test sessions for this cat was 76 degrees; its final threshold (the average of the last ten sessions) was 18 degrees.

The third trained cat (#176) began testing at age six months and required two months to train to the procedure. It was tested over a four month period, then underwent electrophysiological recording. Figure 4c shows the change over time of the threshold for this animal. The average threshold over the first ten test sessions for this cat was 66 degrees. Its final threshold (again, the average of the final ten test sessions) was 33 degrees.

A paired two-group one-tail t-test comparing the average threshold at the beginning of testing (averaged over ten consecutive testing days) to the final threshold for all three cats was performed, with the test hypothesis being that the latter threshold was smaller. The difference was significant at $p=0.01$.

A comparison of the three cats' performance over time is shown in Figure 4d. Note that the thresholds for the first trained cat (#218) were higher overall than the thresholds for the other two cats.

Two other cats (#122 and 40764) were unable to learn the discrimination procedure; they were able to respond to the stimulus on the basis of location of the sound, i.e. they were able to respond when two lateralized speakers were used. However, they were unable to make the switch to discriminating the stimuli coming from one overhead speaker, based on the quality of the sounds. It is interesting to note that these two cats, in contrast to the three successfully trained animals, were generally nervous, timid, and slow-moving. They also underwent electrophysiological recording, and were considered "exposed" control animals; i.e. they had been exposed to the stimuli, but had not performed operant discriminations based upon a perception of differences in these stimuli.

Three out of five cats were able to learn a two-alternative forced choice auditory discrimination procedure. The time required to train to the procedure varied from 3 weeks to 2 months. The threshold was lower for all three cats at the end of testing, as compared to the threshold earlier in training. Performance improvements were continuous and progressive.
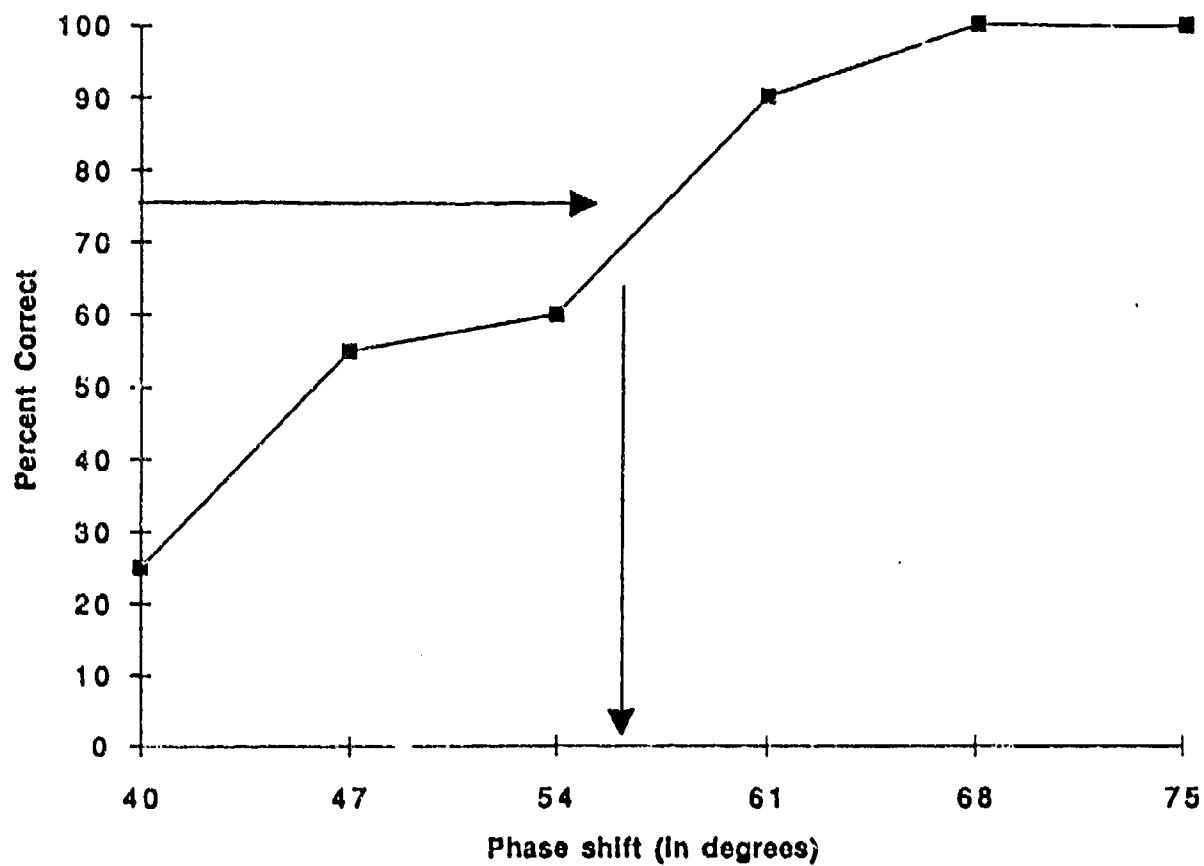
Figure 3. Typical psychophysical function for cat #218.
Threshold envelope phase shift (75% correct) is approximately
57 degrees.
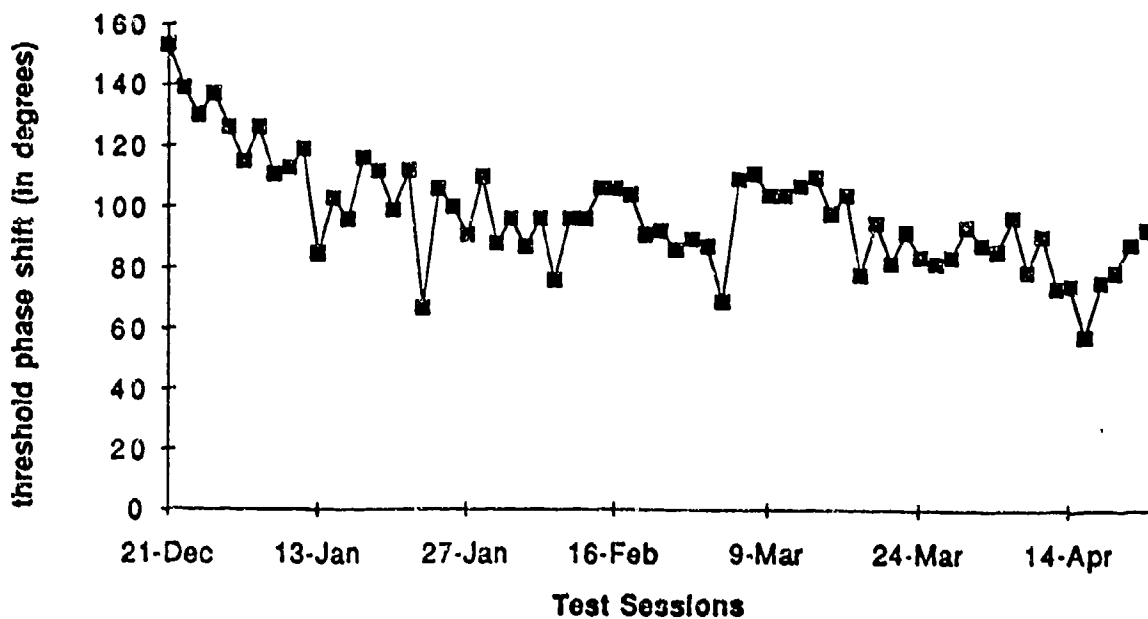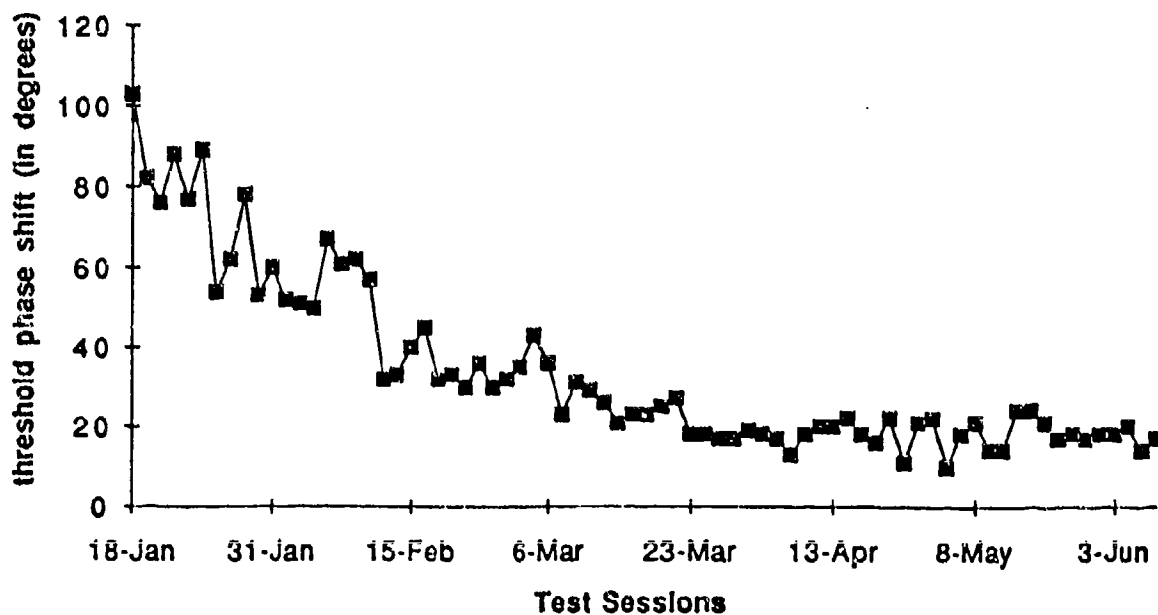
a) Cat #218

b) Cat #293

Figure 4 a) Performance over time of Cat #218, b)
Performance over time of Cat #293.
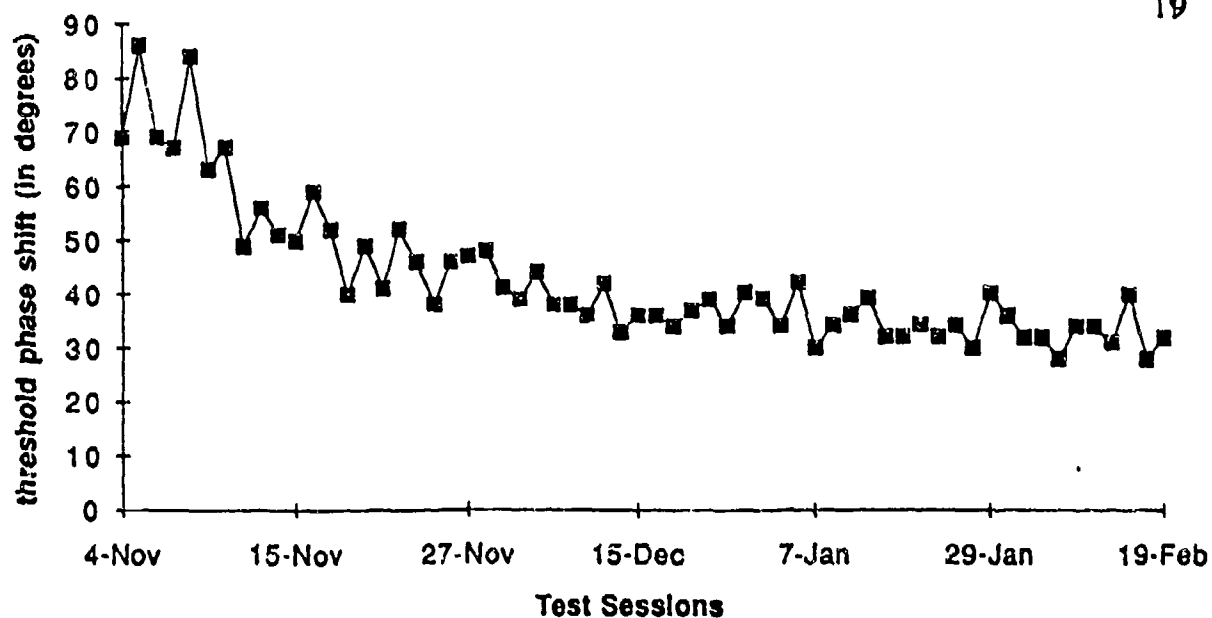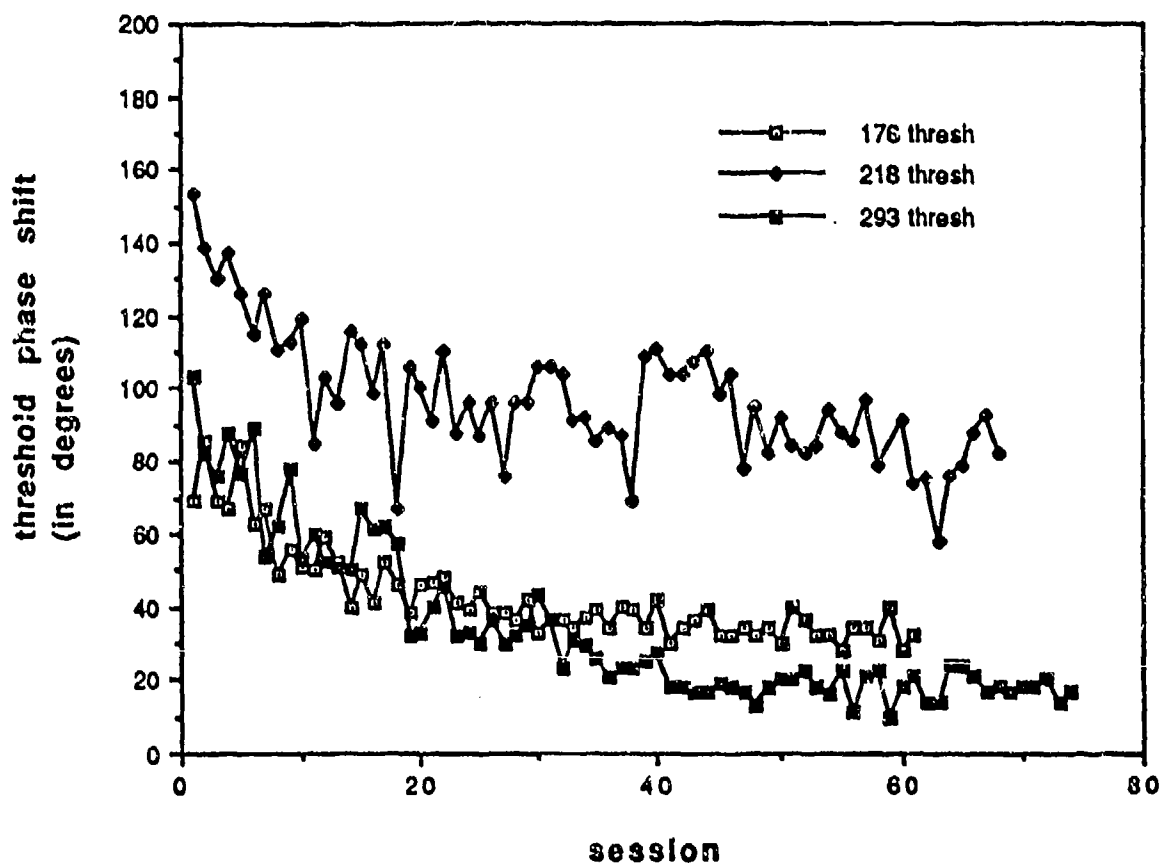
Figure 4 c) Performance over time of Cat #176, d) Comparison
of thresholds over time of three trained cats.

DISCUSSION & CONCLUSIONS

The results of this research are related to various issues, including: a) language acquisition and language learning, and b) what general mammalian auditory processing capacities have to tell us about the evolution of linguistic capabilities.

Language acquisition and language learning

Infant studies have shown that the young infant has a very sophisticated ability to perceive the basic sounds of language (Eimas & Tartter, 1979). Infants even below the age of 4 months are able to make discriminations among steady-state vowels. Trehub (1973) showed that [a] could be distinguished from [i] whether in isolation or in the context of a preceding stop consonant. Infants could also distinguish [i] from [u]. Swoboda, Morse, & Leavitt (1976) showed that the vowel contrast [i] vs [I], as in "beet "vs "bit" was discriminable by infants. Trehub, Endman, & Thorpe (1990) presented complex stimuli that differed in spectral structure to 7 to 8.5 month-old infants. Infants successfully differentiated two spectral structures, even when they varied in fundamental frequency, intensity, or duration. It was concluded that infants classify tonal stimuli on the basis of timbre. Similarly, Clarkson, Clifton, & Perris (1988) presented sounds that differed in their spectral envelopes to 7-month-old infants, who successfully discriminated the sounds. They point out that vowels have multidimensional qualities which distinguish them individually; they argue that timbre is a multidimensional attribute, and point to evidence that shows that "adults' perception of tonal complexes reveals that the spectral envelope is the most important determinant of timbre (Plomp, 1970)" (p.20).

Thus, the ability to process vowel differences based upon changes in spectral envelope seems to be an innate ability in humans. From the animal behavioral data presented, discriminability of changes in spectral envelopes is also present in cats. It may be a basic mammalian auditory processing capacity.

Even though humans seem to have innate speech-processing capabilities, experience does seem to have an effect. Streeter (1976) investigated the discrimination of voicing onset by infants reared in a linguistic environment in which the English voiced/voiceless distinction was not used. The results suggested that "previous exposure to the prevoiced/voiced distinction in the language or perhaps exposure to the acoustic cues underlying this distinction is a prerequisite for discrimination of the contrast. On the other hand, the voiced/voiceless discrimination can be made in the absence of relevant linguistic exposure, since Kikuyu infants do not hear this voicing distinction used in their language. Thus, the results suggest that there is an interaction between nature and nurture; some phonetic or acoustic discriminations may be universal, whereas others seem to require previous relevant exposure" (p.40).

Kuhl, Williams, Lacerda, Stevens, & Lindblom (1992) observed that exposure to a specific language in the first six months of life alters infants' phonetic perception. Infants in America and Sweden were tested with both native- and foreign-language vowel sounds. "Infants from both countries exhibited a language-specific pattern of phonetic perception" (p.606). As pointed out, this perceptual organizing occurs well before the age at which children begin to acquire word meanings, i.e. well before the acquisition of language. It was suggested that "linguistic experience shrinks the perceptual distance around a native-language prototype, in relation to a

nonprototype, causing the prototype to perceptually assimilate similar sounds" (p.608). Thus phonetic prototypes were conjectured to be "fundamental perceptual-cognitive building blocks" (p.608).

The data in this study support the idea that certain basic auditory processing capabilities are present in cats, and that exposure to or experience with the stimuli refines the processing capability and the cortical representation of that stimulus. In the same way that a perceptual organization was set up in the brains of the six-month-old infants, so categorization and learning occur during the acquisition of a first or second language. Learning occurred in these animals as they performed the auditory discrimination, and their brain representations of these behaviorally important stimuli were changed in parallel. It seems plausible that, at birth, certain basic auditory perceptual capabilities are present, and that with linguistic exposure, certain contrasts are sharpened and certain other unused phonetic contrasts drop out of the discriminative repertoire. The cortical organization thus dynamically represents those stimuli to which the human/animal has been exposed and which are informationally and biologically relevant to that creature.

Evolution of linguistic abilities

Clearly humans are more "intelligent" than cats. (If in doubt, try training a cat on an operant discrimination procedure.) Cats were able to learn the paradigm, however, and showed good resolution abilities (at least two out of three cats did), but there was variability in performance, and two other cats were unable to learn the discrimination procedure. Thresholds for the two adept cats were about 3 - 4 times higher than humans for the equivalent ripple density (Keeling, Schreiner, & Jenkins, 1992). Their intelligence as well as their perceptual discrimination abilities

for detecting envelope phase shifts are at a lower level but are on the same continuum as humans.

Watson (1991, ̇s recently reported significant relationships between simple sensory, cognitive, and motor abilities and psychometric intelligence. She reminds us that Spearman (1904) and others reported relationships between pitch discrimination and intelligence. The suggestion is that working memory is a limited-capacity system, and that therefore individuals who can process information rapidly will have more working memory available to process new information. Thus, the sensory ability to perceive differences and make discriminations can be considered part of a processing capacity. Learning calls upon this information-processing ability.

The innate abilities with which humans are born interact with experiential input, especially during development but later as well, to produce a cerebral organization which reflects that interaction. The huge amount of neural modification that occurs within the brain during development allows us to "make sense of the world". As Edelman (1987) says: "it is difficult to imagine the world as it is presented to a newborn organism...the environment presented [to such an organism] is inherently ambiguous; even to animals eventually capable of speech such as ourselves, the world is initially an unlabeled place" (p.3). Adaptive behavior requires initial categorization of salient aspects of the environment so that learning can occur (p.4). Thus, perceptual abilities are very important in order to form concepts. All animals are born with certain basic sensory and discriminative capacities; all must interact successfully with their environment in order to survive. This capacity for learning and for change has allowed Homo sapiens to evolve. In turn, Homo sapiens has

developed speech and language, culture and community. Our capacity for speech has advanced our development of thought, which has changed the world. These faculties have evolved over many eons of time due to the mechanism of neural adaptation. How this cortical change occurs is unknown. Evidence that it occurs is amazing in itself.

ACKNOWLEDGMENTS

REFERENCES

Baru, A.V. Discrimination of synthesized vowels /a/ and /i/ with varying parameters (fundamental frequency, intensity, duration and number of formants) in dog. In G. Fant & M.A.A. Tatham (Eds.), *Auditory analysis and perception of speech*. London: Academic Press, 1975.

Brown, K.A., Buchwald, J.S., Johnson, J.R., & Mikolich, D.J. Vocalization in the cat and kitten. *Developmental Psychobiology*, 1978, 11, 6, 559-570.

Cheney, D.L. & Seyfarth, R.M. Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal behavior*, 1988, 36, 477-86.

Chomsky, N. *Language and Mind*, (2nd ed.). New York: Harcourt Brace Janovich, 1972.

Clarkson, M.G., Clifton, R.K, & Perris, E.E. Infant timbre perception: discrimination of spectral envelopes. *Perception & Psychophysics*, 1988, 43, 15-20.

Dewson, J.H. Speech sounds discrimination by cats. *Science*, 1964, 144, 555-6.

Edelman, G.M. *Neural Darwinism - the theory of neuronal group selection*. New York: Basic Books, 1987.

Eimas, P.D. & Tartter, V.C. On the development of speech perception: mechanisms and analogies. In *Advances in Child Development and Behavior*, H.W. Reese & L.P. Lipsitt, (Eds.), New York: Academic Press, 1979.

Gardner, R.A., & Gardner, B.T. Teaching sign language to a chimpanzee. *Science*, 1969, 165, 664-72.

Heffner, R.S., & Heffner, H.E. Hearing range of the domestic cat. *Hearing research*, 1985, 19, 85-8.

Keeling, D., Schreiner, C.E., & Jenkins, W.M. Discrimination of formant shifts in vowel-like ripple spectra. *Association for Research in Otolaryngology Abstracts*, 1992, 201.

Kuhl, P.K., & Miller, J.D. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science*, 1975, 190, 69-72.

Kuhl, P.K., & Padden, D.M. Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception and psychophysics*, 1982, 32, 542-50.

Kuhl, P.K., & Padden, D.M. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the acoustical society of America*, 1983, 73, 1003-10.

Kuhl, P.K. Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 1992, Jan.31, 255, 606-608.

Liberman, A.M., Cooper, F.S., Shankweiler, D.S., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-61.

Lieberman, P. *The biology and evolution of language*. Cambridge: Harvard University Press, 1984.

Miller, J.D. Audibility curve of the chinchilla. *Journal of the acoustical society of America*, 1970, 48, 2, 2, 513-23.

Miller, J.D. Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America*, 1989, May, 85(5), 2114-2133.

Morse, P.A., & Snowdon, C.T. An investigation of categorical speech discrimination by rhesus monkeys. *Perception and psychophysics*, 1975, 17, 1, 9-16.

Patterson, F. The gestures of a gorilla: Language acquisition in another Pongid. *Brain and Language*, 1978a, 5, 72-97.

Patterson, F. Conversations with a gorilla. *National Geographic*, 1978b, 154, 438-65.

Peterson, G.E., & Barney, H.L. Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 1952, 24, 175.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G.F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.

Premack, D. Language in chimpanzees. *Science*, 1971, 172, 808-22.

Rumbaugh, D.M., Gill, T.V., & von Glaserfield, E.C. Reading and sentence completion by a chimpanzee (Pan). *Science*, 1973, 182, 731-3.

Seyfarth, R., Cheney, D., & Marler, P. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, 1980, 210, 14 Nov., 801-3.

Shipley, C., Carterette, E.C., & Buchwald, J.S. The acoustic phonetics of the cat's meow. Personal communication. Accepted *Journal of the Acoustical Society of America*.

Sinnott, J.M., & Adams, F.S. Differences in human and monkey sensitivity to acoustic cues underlying voicing contrasts. *Journal of the acoustical society of America*, 1987, 82, 5, 1539-47.

Sinnott, J.M., Beecher, M.D., Moody, D.B., & Stebbins, W.S. Speech sound discrimination by monkeys and humans. *Journal of the acoustical society of America*, 1976, 60, 3, 687-95.

Spearman, C. "General intelligence," objectively determined and measured. *American Journal of Psychology*, 1904, 15, 201-293.

Streeter, L.A. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature (London)*, 1976, 259, 38-41.

Suga, N. Auditory neuroethology and speech processing: complex-sound processing by combination-sensitive neurons. In *Auditory function - neurobiological bases of hearing*, G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.). New York: John Wiley & Sons, 1988.

Swoboda, P.J., Morse, P.A., & Leavitt, L. Continuous vowel discrimination in normal and at risk infants. *Child Development*, 1976, 47, 459-465.

Tartter, V. *Language Processes*. New York: Holt, Rinehart & Winston, 1986.

Terrace, H.S., Petitto, L.A., Sanders, R.J., & Bever, T.G. Can an ape create a sentence? *Science*, 1979, 206, 891-902.

Thompson, M., Porter, B., O'Bryan, J., Heffner, H.E., and Heffner, R.S. A syringe-pump food-paste dispenser. *Behavior Research Methods, Instruments, & Computers*, 1990, 22(50), 449-450.

Trehub, S.E. Infants' sensitivity to vowel and tonal contrasts. *Developmental psychology*, 1973, 9, 91-96.

Trehub, S.E., Endman, M.W., & Thorpe, L.A. Infants' perception of timbre: classification of complex tones by spectral structure. *Journal of Experimental Child Psychology*, 1990, 49, 300-313.

Waters, R.S., & Wilson, W.A. Speech perception by rhesus monkeys: the voicing distinction in synthesized labial and velar stop consonants. *Perception and psychophysics*, 1976, 19, 4, 285-9.

Watson, B.U. Some relationships between intelligence and auditory discrimination. *Journal of Speech and Hearing Research*, 1991, June, 34, 621-627.

Discrimination of formant shifts in vowel-like ripple spectra in humans

M. Diane Keeling [a], Christoph. E. Schreiner, and William. M. Jenkins,
Coleman Memorial Laboratory, W.M. Keck Center for Integrative
Neuroscience, University of California, San Francisco; San Francisco, CA
94143-0732.

[a] now at Psychology Dept., Box 5050, University of New Brunswick, Saint
John, Saint John, N.B. E2K 2E2, CANADA

Shortened Title: **Discrimination of formant shifts**

Abstract

The perception of phase shifts in the spectral envelope of vowel-like ripple
stimuli was investigated by measuring the ability of humans to detect frequency
shifts in formant/ripple peak location. The standard ripple stimulus was a
harmonic complex with a fundamental frequency of approximately 45 Hz and a
bandwidth of three octaves centered at 1.38 kHz; the ripple phase of the
standard stimulus was 0 degrees; the ripple depth was 20 dB. The component
amplitudes of the unmodulated signal decreased by 6 dB/octave. Six humans
were tested in a modified 2-Alternative Forced Choice paradigm for their
ability to discriminate envelope phase shifts in these vowel-like ripple stimuli.
Ripple densities from 0.5 ripples/octave to 8 ripples/octave were tested.
Difference thresholds (75% correct) were reported in degrees of envelope
phase shift relative to the standard. The smallest difference detectable in ripple
phase shifts was in the range from 3 to 30 degrees. The highest thresholds were
found for ripple densities between 3.5 and 5 ripples/octave. Thresholds were
lowest for ripple densities below approximately 4 ripples/octave - the peak
spacings that correspond to the majority of formant ratios in English vowels.
PACS 43.71

INTRODUCTION

Vowels are characterized by the positional spacing of formants - component frequency regions which are elevated in amplitude relative to other component frequencies and which are produced by the natural resonances of the vocal tract invoked during production of the vowel. Formant values show considerable variability across speakers however, and even overlap for some vowels (Peterson & Barney, 1952). Nevertheless, listeners are quite proficient at identifying vowels.

How vowels are perceived has been the subject of much speculation and examination. Green (1988) and colleagues investigated the discrimination of a change in spectral shape; they called this mechanism profile analysis; and felt that a critical feature in the detection process is the simultaneous comparison of the intensity level at different frequency regions of the spectrum. Green points out that changes in the relative level of different parts of the acoustic spectrum are perceived by listeners as changes in "sound quality". Sound quality differences exemplify vowel discrimination.

Miller (1989) reports that in the late 1800's Richard John Lloyd formulated the formant-ratio theory of the vowel stating that "vowel quality depends on the intervals between the resonances, not their absolute values" (p.2155). Miller's auditory-perceptual interpretation of the vowel is a descendant of the formant-ratio theory and proposes that changes in spectral *patterns* of formants effectively categorize vowel sounds.

The ratios of the formants may be an important cue. We calculated the ratios of the first and second, and second and third formants; since however, frequency is represented on a logarithmic scale along the basilar membrane within the cochlea, we computed the ratios of the formants in

units per octave (Miller,1989). The spacing of the formants or ripple density for the formant ratio F1/F2 and F2/F3 of 10 English vowels is shown in Figure 1 with values grouped for man, woman, and child. The average frequency values for the formants were taken from .... (see The Science of Sound). Ripple densities extend from 0.3 ripples/octave to 4.5 ripples/octave, with the majority being less than 3 ripples/octave. Note that the ripple densities for most vowels across men, women, and children were quite similar.

We hypothesized that the perception of variations in the the position of spectral peaks is dependent on the spacing of the peaks. We set out to measure humans' ability to discriminate among vowel-like complex stimuli. We used a) a broad-band stimulus (3 octaves wide) in order to simulate a natural vowel, b) a harmonic spectrum, again in order to mimic vowels, c) a sinusoidal envelope pattern on a logarithmic scale, and we chose d) 20 dB as the depth of modulation of the standard stimuli. The precision with which the peak shifts could be determined, and the variation of this precision across several ripple densities was expected to shed light upon the importance of small changes in formant location for vowel processing, and thus elucidate the perception of vowels.

METHODS

Stimulus generation and presentation and response tabulation were controlled by a Macintosh IIx computer. The LabVIEW application (National Instruments Corporation) was used to automate procedures for data acquisition and instrument control. A program was created using a DSP board and output via a 16 bit Audio board, that allowed the presentation of ripple stimuli to the subject in the sound booth, and to record and tabulate on-line the responses of the subject. An input/output

board served as the interface between the computer and the subject. A Crown D-75 amplifier amplified the signals.

Subjects and Procedure

Six persons (two men and four women; mean age: 35 years, range 22 to 44 years) with normal hearing in the tested frequency range as determined by professional audiological assessment, were tested using a modified 2-Alternative Forced Choice (2-AFC) procedure. The subject was seated in a sound-attenuating chamber (IAC). The subject was told that upon initiating a trial by a center button press on a panel held on the lap, s/he would hear through headphones (AKG -K240DF) three sounds. The stimulus presentation was monaural, to the left ear. The task of the subject was to indicate which of the three sounds was different from the other two. The first stimulus served as a standard stimulus and was always the same for a single test session. Either the second or the third stimulus was different from the other two; the subject was to indicate the odd stimulus by a button press: left button if the second stimulus was different from the other two, right button if the third stimulus was different. A green light was illuminated if the response was correct. In each test session there were 147 trials (21 different triads, each presented 7 times). A session usually took about ten minutes. Subjects usually completed three or four sessions per sitting.

Stimulus

The formant values for ten English vowels (values takes from Peterson & Barney, 1952) were converted into ratios for the first formant and second formant (F1/F2) and for the second formant and third formant (F2/F3). These values were then converted to a logarithmic scale, giving the number of peaks or ripples per octave, also termed the ripple density.

Figure 1 shows the ripple densities for F1/F2 and for F2/F3. Values range from 0.3 to 4.5 ripples per octave, with the majority below 3 ripples/octave.

In order to test the discriminability of envelope phase shifts as a function of formant spacings, a ripple stimulus was created with the following characteristics: a) the carrier consisted of a harmonic series with a fundamental frequency of approximately 45 Hz (to give closely spaced harmonics) and a 6 dB/octave decline of the component amplitudes; b) the bandwidth of the stimulus was 3 octaves; c) the spectral envelope of the signal was represented by a sinusoid on a logarithmically scaled frequency axis, the frequency of the envelope sinusoid was referred to as ripple density and measured as ripples/octave, varied from 0.5 to 8 ripples/octave; and d) the modulation depth of the envelope (ripple depth) was linear on a dB scale, and was set at 20 dB.

The standard ripple stimulus (Figure 2) always contained a central formant that was centered at 1.38 kHz. This frequency is approximately centered within the range of formant frequencies for English vowels: (270 Hz to 3010 Hz; Peterson & Barney, 1952). The stimulus, being three octaves wide, extended from 340 Hz to 2,672 Hz. Within each trial, it was deemed necessary that the total energy content of each stimulus be kept constant; this was done by varying the phase of the envelope of the stimulus, instead of the ripple density.

The ripple phase of the standard stimulus was 0 degrees, i.e. the signal (or odd) ripple stimulus had a phase varying from 1 degree to 180 degrees. ----Figure 3 illustrates a ripple stimulus, with ripple density of 2, shifted 45 degrees. As can be seen, the frequency values of the peaks are shifted. The phase was shifted in both a positive and a negative direction

relative to the standard stimulus; for the positive phase shift, frequency peaks shifted to lower frequencies, for the negative phase shift, frequency peaks shifted to higher frequencies.

The intensity of the stimulus was 60 dB SPL. All three stimuli within each trial were presented at the same level, as it was found to be too difficult for subjects to ignore intensity variations of more than 2-3 dB while discriminating phase shifts. However, the intensities of the stimuli varied randomly by +/- 3 dB around the standard 60 dB SPL stimulus across trials.

Threshold computation

The minimum phase shift of the envelope that could be detected with this modified 2-AFC method was computed in the following way. In each testing session, one-hundred forty-seven trials were presented. In each trial, either the second or third stimulus was different from the other two. The subject was forced to make a choice and was thus either correct or incorrect on each trial. At the beginning of the session, the range of phase shift values was set, with 1 to 180 degrees being the largest range. The ideal was to present a range of values that challenged the subject in order to define the true threshold, but that was not so difficult as to produce extreme frustration ----for the subject (Green, 1990). Twenty-one values within this range were then randomly presented as the "different stimulus", each being presented 7 times (thus equaling a total of 147 trials). The number of responses correct at each phase shift value were tabulated on-line by computer. On screen could be seen the overall percent of responses correct. In addition, the percent correct on the left side (left button, indicating second stimulus different) and the percent correct on the right side (right button, indicating third stimulus different) were shown, in order to detect response bias for either stimulus. At the end of the test session, stimulus and response values were stored, and the thresholds computed. A not discriminated stimulus resulted in a correct response percentage of 50%. Discrimination thresholds were considered as 75% correct, and were calculated by a linear regression formula applied to the psychometric function. The threshold was thus the minimum phase shift, in degrees, detectable at 75% correct. Ripple densities from 0.5 ripples/octave to 8 ripples/octave were tested. Each ripple density was tested at least twice, totalling 294 presentations. A threshold was obtained for each subject at

each ripple density value. Further conditions were presented as described below.

a. Ramped vs. Unramped Envelope Condition

In order to evaluate whether the subjects used spectral envelope information across the entire width of the spectrum or were biased by changes at the edges of the spectrum, a ripple stimulus was created in which the spectral envelope was ramped at the low- and high-frequency end of the signal. In the ramped portions, the depth of modulation changed gradually over 0.5 octaves from a modulation depth of 0 dB at the high and low-frequency ends of the spectrum to the full value of 20 dB. Since the total signal width was 3 octaves, the fully modulated envelope was realized over 2 octaves.

Two subjects of the originally described six subjects were tested on the ramped ripple stimulus at 7 different ripple densities, ranging from 0.5 ripples/octave to 6 ripples/octave.

b. Depth of Modulation

The ripple stimulus had been created with the possibility of varying several of its parameters. One such parameter which was varied for one set of tests was depth of modulation of the envelope of the ripple stimulus. It was varied across the values 2.5, 5, 7, 10, 20, and 30 dB, for each of the ripple densities of 0.5, 1, 2, and 4 ripples/octave. Two subjects were tested.

c. Intensity

The intensity of the stimulus had been 60 dB SPL throughout all previous testing. For this set of tests, the overall intensity levels were varied across 30, 40, 50, 60, and 70 dB SPL. Thresholds were obtained at

3 different ripple densities (one, three and five ripples/octave) for three subjects.

RESULTS

Thresholds for the just detectable shift in envelope phase were obtained for each subject at various ripple densities. The threshold represents the smallest envelope phase shift in degrees that the subject can detect with 75% or higher accuracy. Figure 4a shows a typical psychometric function for a representative subject. Performance (percent correct) is plotted as a function of degrees shifted negative and positive, in this case for a ripple density of 1 ripple/octave. Figure 4b shows the performance of the same subject for a ripple density of 4 ripples/octave. In each test session, each value was presented 7 times. Each ripple density session was presented at least twice. Values shown represent the average performance over 14 stimulus presentations. The thresholds for a ripple density of 1 ripple/octave were +2.3 for positive phase shifts and -4.7 degrees for negative phase shifts. The thresholds for a ripple density of 4 ripples/octave were +13.5 and -22.5 degrees, respectively.

Figures 5 and 6 illustrate the performances for each of the six subjects, presenting thresholds (minimum phase shift detectable at 75% correct) in firstly positive, then negative directions, as a function of ripple density. Note that the overall dependence of the phase threshold is nonmonotonic, and that for positive phase shifts, the threshold generally increased with an increase in ripple density to a maximum at about 4 to 5 ripples/octave, then generally decreased. For negative phase shifts, similarly nonmonotonic behavior was seen but there was greater variability in the shapes of the functions, and thresholds were highest at anywhere from 3 to 8 ripples/octave.

Keeling, M.D. JASA

Figure 7 presents thresholds across all tested ripple densities pooled for all subjects: a) for positive phase shifts and b) for negative phase shifts. For positive phase shifts, thresholds were highest at 4.5 and 5 ripples/octave. The lowest threshold was at 1.5 ripples/octave, where variability among subjects was lowest. For the negative phase shifts, the highest thresh. 1 was at a ripple density of 4 ripples/octave. The lowest threshold was again at 1.5 ripples/octave, where variability among subjects was also lowest. Tables I and II present the means and standard deviations for each of the averaged values across all ripple densities, for envelopes shifted in both positive and negative directions. A repeated-measures ANOVA was carried out for both positive threshold and negative threshold data. For the positive data, the overall F for threshold differences between different ripple densities was significant at p=0.03. Comparison among individual means using the Scheffe test for the various ripple densities showed differences significant at p<0.05 for the following ripple density pairs: .5 vs. 4, 4.5, and 5; 1 vs. 4, 4.5, and 5; 1.5 vs. 4, 4.5, and 5; 2 vs. 4.5 and 5; 2.5 vs. 4, 4.5, and 5; 3.5 vs. 4.5, and 5; 4.5 vs. 8; and 5 vs. 8. For the negative threshold data, the overall F was not significant (p=0.07), although a similar general trend was observed. However, in comparing the peaks for these subjects in the positive vs. the negative phase shift condition (two-group paired t-test), there was a statistically significant (p=0.01) difference between the location of the peaks in the two conditions.

Thus, in general, what was observed for both positive and negative envelope phase shifts was an increase in threshold as ripple density increased up to approximately 4 to 5 ripples/octave, then a decrease, at least to the extent of the ripple densities tested here, i.e. 8 ripples/octave.

a. Ramped vs. Unramped Condition

Keeling, M.D. JASA

In order to evaluate whether the subjects used spectral envelope information across the entire width of the spectrum or were biased by changes at the edges of the spectrum, a ramped ripple stimulus was created. The results of the tests on the ramped ripple stimulus as compared to the original un-ramped stimulus are shown in Figure 8. The positive phase shift thresholds are plotted as a function of different ripple densities for the ramped and un-ramped conditions. For subject (dk), the thresholds were somewhat higher in the ramped condition for all ripple densities; however, the shapes of the functions were very similar, with the highest thresholds at a density of 5 ripples/octave. Comparison of the two conditions (ramped vs. unramped) using a paired t-test revealed that they were different for this subject (p=0.01). For subject (bc), the thresholds were very similar in the ramped and un-ramped conditions. Thresholds were higher in the ramped condition for 5 of 8 ripple densities. Again, the shapes of the functions were very similar, this time with the highest thresholds at 4 ripples/octave. For this subject, when the two conditions were compared using a paired t-test, they did not significantly differ (p=.89). Thus, for each subject the global pattern of threshold distribution was the same. However, since there were differences between some of the means for one of the subjects, there is an indication that different subjects may use different spectral regions in their discrimination processing of envelope shifts.

b. Depth of Modulation

The depth of modulation was varied, to shed light on the contrast required to discriminate spectral energy peaks. The results from these test sessions are shown in Figure 9. The positive phase shift thresholds are plotted as a function of the depth of modulation, for four different ripple

densities. For both subjects, similar trends were observed, namely a) a rapid change of threshold with modulation depths between 2.5 and 7.5 dB, and b) a relatively shallow threshold curve for modulation depths of 10, 20 and 30 dB. In both subjects, thresholds were highest at a depth of modulation value of 2.5 dB for all four ripple densities tested. Thresholds then rapidly decreased at depth of modulation values of 5, and 7.5 dB. For the first subject (dk), there was a slight increase at a depth-of-modulation value of 10 dB across all 4 ripple densities, and then generally, a decline in threshold. The absolute lowest threshold for this subject was at a depth of modulation of 20 dB with ripple densities of 1 and 0.5 ripples/octave. For subject bc, the threshold decreased from 2.5 to 10 dB depth of modulation for all ripple densities except one (ripple density = 4 ripples/octave), where it was increased. The absolute lowest threshold for this subject was at a depth of modulation of 30 dB with a ripple density of 0.5 ripples/octave. For both subjects, lowest thresholds were at either 20 or 30 dB depth of modulation for all ripple densities.

It appeared from the results of Figure 9 that the thresholds at the lowest three depths of modulation: (2.5, 5, and 7.5 dB) grouped separately from the thresholds derived for the second three depths of modulation: (10, 20, and 30 dB). Therefore, threshold values were collapsed across ripple densities and divided into these two groups, and compared. A linear regression analysis was performed. The values for the two subjects were combined, i.e. at each point, they were added and averaged. The regression for the first group of modulation depths was compared to the regression for the second group; the regression coefficients (slopes) were significantly different at $p < 0.01$. Figure 10 illustrates the comparison of

the regression lines for the two groups of depth of modulation for the two subjects' combined values.

Thus, it appears that below a depth of modulation value of approximately 7.5 dB, the threshold for detecting ripple stimulus envelope phase shifts increases dramatically. Above the depth of modulation value of 7.5 to 10 dB, the threshold varies only slightly , at least for the depths of modulation presented herein, namely, 10, 20, and 30 dB.

c. Intensity

Vowels are spoken and differentiated over a wide range of levels. Therefore, it is important to know whether the ripple phase thresholds measured are intensity independent, or whether they vary according to the intensity at which these stimuli are presented. Results are shown in Figure 11. The positive phase shift thresholds are plotted as a function of ripple densities for five different intensity levels.

For subject kk, mean and standard deviation for the variability due to intensity at each ripple density, are shown in Table III. When an ANOVA was computed comparing the values for three ripple densities across five intensities, the overall F was significant at $p=0.01$, but there were no significant differences among the various intensities (Fisher PLSD, Scheffe, and Dunnett).

For subject dk, similar results were found. The mean and standard deviation for the variability due to intensity at each ripple density is shown in Table IV. When an ANOVA was computed comparing the values for three ripple densities across five intensities, the overall F was significant at $p<0.01$ (the curve was not flat), but there were no significant differences among the various intensities (for all post-hoc tests: Fisher PLSD, Scheffe, and Dunnett).

For subject bc, results were again similar. The mean and standard deviation for the variability due to intensity at each ripple density, is shown in Table V. When an ANOVA was computed comparing the values for three ripple densities across five intensities, the overall F was not significant (p=0.10); there were no significant differences among the various intensities. Thus in general threshold does not appear to be strongly dependent on intensity, at least over the tested range.

Phase shifts compared to frequency differences

When the spectral envelope of the ripple stimulus is shifted, the frequency values at the location of the peaks change. Actually the amplitude at the various component frequencies changes so that different peaks occur but the frequencies of the components are always constant. Nevertheless, it is possible to calculate the frequency difference that would correspond to a particular ripple stimulus envelope phase shift by using the formula: phase shift = (log (b/a))/log 2)*360* ripple density, where 'a' is the center frequency of the standard stimulus, in this case, 1.38 kHz, and 'b' is the center frequency of the comparison stimulus. We compared the frequency difference equivalent to the mean threshold envelope phase shift to frequency difference limens previously reported for pure tones. The frequency difference (in Hz) at the center frequency that corresponds to the average threshold for positive envelope phase shifts at each ripple density is listed in Table VI. Similar frequency differences for negative envelope phase shifts are listed in Table VII.

Wier, Jesteadt, and Green (1977) reported frequency difference limens (DLs) for humans, incorporating their own and previous data (see p.30). Fay (1988) reports that the frequency discrimination threshold (or

frequency difference limen) at a frequency of 1000 Hz is 1.9 Hz and at 2000 Hz is 3.2 Hz for humans (p.458). Extrapolating from this information, one may estimate the frequency dl at 1.38 kHz to be approximately 2.5 Hz. From Tables VI and VII, we observe that the smallest shift detected was equivalent to a frequency shift of 3.89 Hz (for a positive shift at 8 ripples/octave). The negative shift at 8 ripples/octave was quite similar (3.96 Hz). Thus, the smallest discriminable frequency difference observed is larger than the frequency DL value reported by Fay, by almost a factor of 2. Although pure tone discrimination and ripple discrimination are clearly not the same, the comparison may elucidate psychophysical processing mechanisms.

DISCUSSION

Humans appear to be most adept at discriminating ripple envelope shifts at spacings of 3.5 or fewer peaks/octave. The spacings of English vowels are mostly of 3 or less peaks/octave. Not surprisingly, perhaps, humans are best at discriminating changes within the spacings of English vowels. Since formants are determined by natural resonances of the human vocal tract, the implication arises that the human auditory system has either evolved to be best at discriminating those parameters that are important in speech, and/or through adaptive learning, has developed a formal representation of this most heavily practiced peak-spacing domain..

Flanagan (1955) first provided detailed measures of difference limens (DLs) for vowel formant frequencies. In his studies, for each synthesized vowel three out of four of the formants were held constant, while either F1 or F2 was varied. In the first set of tests, F2 was 1500 Hz, F3 measured 2500 Hz, F4 was 3550 Hz, and F1 was set at 300, 500, or 700

Hz. Frequency variations used in determining the DLs for F1 were +/-10, +/-20, +/-30, +/-40, +/-50, +/-60, and +/-70 Hz.

In a second set of tests, F2 was set at 1000, 1500, or 2000 Hz; F1 was held constant at 500 Hz; the frequency variations used in determining the DLs for F2 were +/-25, +/-50, +/-75, +/-100, +/-125, +/-150, and +/-175 Hz. Subjects heard two stimuli and were asked to report whether the second sound was the same as or different from the first. The 50% point of stimuli judged different was taken as the DL. The mean just-discernible frequency shift was approximately 3-5% of the formant frequency.

Comparing the Flanagan results to the present ripple study results is complicated by the fact that the peak or formant spacings in synthesized vowels were not identical among all peaks within a vowel, as they were in our ripple study. However, ignoring for the moment the formants above F2, the various spacings between F1 and F2 studied may be considered. In the first set of tests in which F1 was varied, the frequency values of F1 and F2 were i) 300 Hz and 1500 Hz; ii) 500 Hz and 1500 Hz; and iii) 700 Hz and 1500 Hz, respectively. In the second set of tests, in which F2 was varied, the frequency values of F1 and F2 were i) 500 Hz and 1000 Hz; ii) 500 Hz and 1500 Hz; and iii) 500 Hz and 2000 Hz. The equivalent ripple densities are 0.4, 0.6, 0.9, 1, 0.6, and 0.5 ripples or peaks/octave.

Flanagan reported that human subjects could detect a 3 to 5% shift in spectral peak position. That would be equivalent to a 55 Hz shift at the center frequency of 1.38 kHz in the present human ripple envelope phase shift study. The equivalent envelope phase shift for a ripple density of 1 ripple/octave, is 20.3 degrees. For a ripple density of 0.5, the phase shift value equivalent to 55 Hz is 10.1 degrees. As will be recalled from Tables VI & VII, human phase shift thresholds, at a ripple density of 1

ripple/octave, were 7 degrees for positive envelope phase shifts and 8.6 degrees for negative envelope phase shifts, averaging 7.8 degrees. The human phase shift thresholds at a ripple density of 0.5 were 7.3 degrees for positive envelope phase shifts, and 13.5 degrees for negative envelope phase shifts, averaging 10.4 degrees. The equivalent frequency DLs were 21 Hz and 54 Hz for ripple densities of 1 and 0.5 ripples/octave. This information is presented in Table VIII.

Thus the frequency DL values observed by Flanagan of approximately 3-5% of the frequency tested matched the measures of the detectable shifts for 0.5 ripples/octave stimuli, but distributions at 1 ripple/octave were poorer for vowel peak shift than would be expected from the current measures of ripple peak shifts in our human study.

Van Veen & Houtgast (1985) have earlier argued that the depth of modulation plays only a minor role in vowel discrimination. They also observed that the ripple densities that are most important for the perception of differences in spectral sharpness are around 2 ripples/octave. In the current study, it was observed that across four ripple densities (0.5, 1, 2, and 4 ripples/octave), depth of modulation changes affected discrimination similarly, with depths below 7.5 dB producing a dramatic increase in threshold. It is difficult to compare these results directly with those of van Veen & Houtgast because they used a different filtering method to spectrally sharpen their stimuli. However, in contradistinction to their general conclusions, the present study reveals that for depths of modulation below about 7.5 dB, shifts in frequency peak positions rapidly come to be far more difficult to detect.

Shamma & Vranic (1993) reported that the detection of change in symmetry in spectral shape was more dependent on peak frequency (pitch

effect) than was the detection of a change in bandwidth. They also reported that "the listeners required less training and exhibited higher sensitivity (lower thresholds) for the detection of symmetry than bandwidth factor peak changes" (p.47). Vranic, Versnel, & Shamma (1993) also presented both single and double spectral peaks, but did not report a covariance of neural responses of AI neurons in ferret cortex with the number of spectral peaks. An expanded report concerning the effect of single and double peaks on psychophysical discrimination of these complex stimuli would be revealing, and directly comparable to the present study. Their data imply that spectral shape is very informative to the nervous system, and that many parameters of complex stimuli provide discriminatory cues.

In Figure 12, the average minimum envelope positive phase shifts detected across the ripple densities tested - 0.5 to 8 ripples/octave - are shown. Also shown are the frequency difference limen (dl) values for 4 different frequency values: 1 kHz, 3 kHz, 6 kHz, and 8 kHz, plotted against ripple density. As for formant ratios, frequency difference limen/frequency ratios can be converted to a logarithmic scale, and termed "peaks" or "events per octave", or "ripple density". Thus, as can be seen, as ripple density increases, the number of events/octave that the frequency difference limen represents also increases. It should be noted that the slope of the envelope phase shift threshold and the slopes of the frequency difference limen functions are generally similar. This suggests that the discriminative functions involved in frequency discrimination and in ripple envelope shift discrimination may be similar. However as was earlier pointed out , the frequency dl for a frequency of 1.38 kHz (the center frequency of the human ripple stimulus) is approximately 2.5 Hz, while the

smallest discriminable phase shift for ripple stimuli observed in this study was equivalent to a frequency value of approximately 4 Hz. Thus, the values do not coincide.

The critical band ratio converted to ripple density is also shown in Figure 12. Its value is approximately 4 ripples/octave, and corresponds to the transition to increased thresholds, suggestive of an interference of the filter (critical band) spacing with the spacing of major spectral peaks in the stimulus. Thus peaks of spacings of less than 4 per octave would span the critical band border, and discriminability of envelope phase shifts within the range from 0.5 to approximately 4 ripples/octave would be good and fairly constant, as it appears to be. Peaks spaced at 4 per octave or more would fall within single critical bands; therefore envelope peak shifts would not be as easily discriminable, and the threshold should increase.

It can be observed that the minimum envelope phase shift function decreases after a peak at about 5 ripples/octave. There is an increase again at 7 ripples/octave and then another decrease at 8 ripples/octave. All subjects showed a decrease in threshold after a peak around 4 or 5 ripples/octave. Some did not show a second later increase in threshold. Why should the threshold decrease after approximately 5 ripples/octave? As mentioned, some subjects showed a second peak in threshold, while others did not. It may be that there is a second interference with peak spacing discrimination related to a function of the critical band.

Studies by Schreiner, Calhoun, & Keeling (1993) have shown that "untrained" cats have the best neural response to ripple densities between 1 and 2 ripples/octave. This implies that there may be something inherently biologically appealing about the octave scale, in that one event per octave, or one event for every doubling of frequency, facilitates discrimination.

With the critical band being approximately one-quarter to one-third octave wide, one might expect that spectral peak spacings of less than 3 or 4 would be fairly easily discriminated, since the peaks fell into different critical bands. It is also the case that most spacings of formants in English vowels are less than 4 peaks per octave. Thus animal electrophysiological data and the human psychophysical data indicate that formants are located within a range of spacing values that coincide with maximally discriminable perceptual information elements.

Infant studies have shown that the young infant has a very sophisticated ability to perceive the basic sounds of language (Eimas & Tartter, 1979). Infants even below the age of 4 months are able to make discriminations among steady-state vowels. Trehub (1973) showed that [a] could be distinguished from [i] whether in isolation or in the context of a preceding stop consonant. Infants could also distinguish [i] from [u]. Swoboda, Morse, & Leavitt (1976) showed that the vowel contrast [i] vs [I], as in "beet "vs "bit" was discriminable by infants. Trehub, Endman, & Thorpe (1990) presented complex stimuli that differed in spectral structure to 7 to 8.5 month-old infants. Infants successfully differentiated two spectral structures, even when they varied in fundamental frequency, intensity, or duration. It was concluded that infants classify tonal stimuli on the basis of timbre. Similarly, Clarkson, Clifton, & Perris (1988) presented sounds that differed in their spectral envelopes to 7-month-old infants, who successfully discriminated the sounds. They point out that vowels have multidimensional qualities which distinguish them individually; they argue that timbre is a multidimensional attribute, and point to evidence that shows that "adults' perception of tonal complexes

reveals that the spectral envelope is the most important determinant of timbre (Plomp, 1970)" (p.20).

Thus, the ability to process vowel differences based upon changes in spectral envelope seems to be an innate ability in humans. From the animal behavioial data (Schreiner et al. 1993), discriminability of changes in spectral envelopes is also present in cats. It may be a basic mammalian auditory processing capacity.

ACKNOWLEDGMENTS

REFERENCES

Clarkson, M.G., Clifton, R.K, and Perris, E.E.(1988). "Infant timbre perception: discrimination of spectral envelopes," Perception & Psychophysics,43, 15-20.

Delattre, P., Liberman, A.M., Cooper, F.S., and Gerstman, L.(1952). "An experimental study of the acoustic determinants of vowel color: observations on one- and two-format vowels synthesized from spectrographic patterns," Word,8, 195-210.

Eimas, P.D. and Tartter, V.C.(1979). "On the development of speech perception: mechanisms and analogies". In Advances in Child Development and Behavior, edited by H.W. Reese & L.P. Lipsitt (Academic Press, New York).

Fay, R.R.(1988). Hearing in Vertebrates: a psychophysics databook. (Hill-Fay Associates, Winnetka, Illinois).

Flanagan, J.L. (1955). "A difference limen for vowe formant frequency," Journal of the Acoustical Society of America, May, 27, 3, 613-617.

Green, D.M. (1988) Profile analysis - Auditory Intensity Discrimination. (Oxford University Press, New York, N.Y.).

Green, D.M. (1990). "Stimulus selection in adaptive psychophysical procedures," Journal of the Acoustical Society of America, June, 87, 6, 2662-2674.

Miller, J.D. (1989). "Auditory-perceptual interpretation of the vowel," Journal of the Acoustical Society of America, 85, 5, 2114-2134.

Peterson, G.E. & Barney, H.L. (1952). "Control methods used in a study of vowels," Journal of the Acoustical Society of America, 24, 175.

Pickles, J.O. (1988). An Introduction of the Physiology of Hearing, 2nd Edition. (Academic Press, London).

Pickles, J.O. (1975). "Normal critical bands in the cat," Acta Oto-laryngol., 80, 245-254.

Plomp, R. (**1970**). "Timbre as a multidimensional attribute of complex tones,". in *Frequency analysis and periodicity detection in hearing*,, edited by R. Plomp and G.F. Smoorenburg (Sijthoff,Leiden).

Schreiner, C.E., Calhoun, B., and Keeling, D. (**1993**). "Physiology and topography of cortical neurons explored with vowel-like ripple-spectra." Association for Research in Otolaryngology Abstract, 161.

Shamma, S.A. and Vranic, S. (**1993**). "Detection of changes in spectral peak shape varies with peak amplitude, spectral density, and pitch," Association for Research in Otolaryngology Abstracts, 187.

Swoboda, P.J., Morse, P.A., and Leavitt, L. (**1976**). "Continuous vowel discrimination in normal and at risk infants,". Child Development, 47, 459-465.

Trehub, S.E. (**1973**). "Infants' sensitivity to vowel and tonal contrasts," Developmental psychology, 9, 91-96.

Trehub, S.E., Endman, M.W., and Thorpe, L.A. (**1990**). "Infants' perception of timbre: classification of complex tones by spectral structure," Journal of Experimental Child Psychology, 49, 300-313.

Van Veen, T.M., and Houtgast, T. (**1985**). "Spectral sharpness and vowel dissimilarity," Journal of the Acoustical Society of America, Feb., 77(2), 628-634.

Vranic, S., Versnel, H., & Shamma, S.A. (**1993**). "Single and double spectral peaks: psychoacoustical and physiological results," Association for Research in Otolaryngology Abstracts, 188, p.47.

Wier, C.C., Jesteadt, W., and Green, D.M. (**1977**). "Frequency discrimination as a function of frequency and sensation level, "Journal of the Acoustical Society of America, Jan., 61, 1, 178-184.

Figure 1. Ripple Density for F1/F2 and for F2/F3. Values range from 0.3 to 4.5 ripples per octave, with the majority below 3 ripples/octave.

Ripple Stimulus
(centered at 1.38 kHz, 3 octaves wide,
ripple density = 1 ripple/octave)
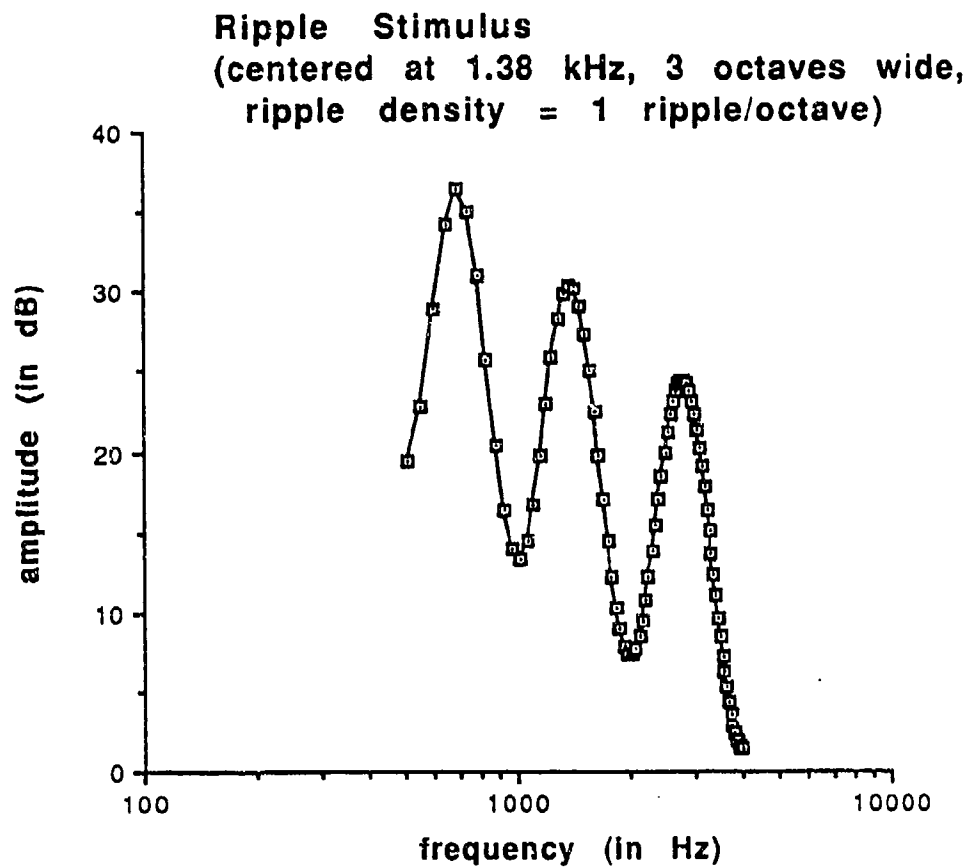
Figure 2. Ripple Stimulus for Human Testing

Figure 3. Phase-shifted ripple stimuli, with ripple density of 2, shifted 45 degrees. As can be seen, the frequency values of the peaks are shifted.

Figure 4. Performance of subject dk on a ripple stimulus of ripple density a) 1 ripple/octave and b) 4 ripples/octave. The thresholds for a ripple density of 1 ripple/octave were +2.3 and -4.7 degrees. The thresholds for a ripple density of 4 ripples/octave were +13.5 and -22.5 degrees.

Figure 5. Performance of subjects across ripple densities, for positive envelope phase shifts.

Figure 6. Performance of subjects across ripple densities, for negative envelope phase shifts.

Keeling, M.D. JASA

Figure 7. Thresholds pooled for all subjects a) for positive
phase shifts and b) for negative phase shifts, showing means and
standard deviations at each ripple density.

Keeling, M.D. JASA

Figure 8. Thresholds for ramped and unramped stimulus
conditions for two subjects. Some subjects may use the edge of
the spectrum in their discrimination processing.

Figure 9. Threshold as a function of depth of modulation (two subjects). Two subjects were tested across four ripple densities at six different depths of modulation.

Figure 10. Comparison of the regression lines for the two groups of depth of modulation for the two subjects' combined values. Below a depth of modulation value of 7.5 dB, the threshold for detecting ripple stimulus envelope phase shifts increases dramatically.

Keeling, M.D. JASA

34

Figure 11. Threshold at various intensities (three subjects).
Threshold does not appear to be strongly dependent on intensity
over the tested range.

Figure 12. Frequency Difference Limen re Envelope Phase
Shift. Human thresholds plotted against ripple density,
frequency difference limens, and critical band density.

ripple density	mean threshold shift	standard deviation
0.5	7.3	4.4
1	7	3.6
1.5	5.9	2.5
2	9	4.2
2.5	8.7	5
3	14.2	5.9
3.5	11	3.9
4	20.5	6.9
4.5	25.5	8.4
5	25.5	10.6
6	13.9	7.1
7	21.1	5.3
8	11.7	4.7

Table I. Mean threshold shift and standard deviation at each ripple density for positive phase shifts (N=6)

ripple density	mean threshold shift	standard deviation
0.5	13.5	3.4
1	8.6	4
1.5	7.8	1.8
2	12.4	5.3
2.5	14.4	5.4
3	18.3	10
3.5	15.1	6.9
4	24.7	5.6
4.5	15.6	4.5
5	14.8	4.8
6	16.6	5
7	15.3	4
8	11.9	7.1

Table II. Mean threshold shift and standard deviation at each ripple density for negative phase shifts (N=6)

	mean threshold	standard deviation
rd=1	7.5	2.4
rd=3	10.7	2.4
rd=5	14.8	3.8

Table III. Means and standard deviations across intensities for subject kk.

	mean threshold	standard deviation
rd=1	3.3	1.5
rd=3	9.5	2.0
rd=5	20.9	4.8

Table IV. Means and standard deviations across intensities for subject dk.

	mean threshold	standard deviation
rd=1	4.5	1.9
rd=3	5.9	3.0
rd=5	9.4	4.6

Table V. Means and standard deviations across intensities for subject bc.

ripple density	mean threshold (in degrees)	difference in frequency (in Hz)
0.5	7.3	-38.25
1	7	-18.47
1.5	5.9	-10.41
2	9	-11.91
2.5	8.7	-9.22
3	14.2	-12.52
3.5	11	-8.33
4	20.5	-13.55
4.5	25.5	-14.97
5	25.5	-13.48
6	13.9	-6.14
7	21.1	-7.99
8	11.73	-3.89

Table VI. Frequency differences corresponding to threshold positive phase shifts.

ripple density	mean threshold (in degrees)	difference in frequency (in Hz)
0.5	13.5	73.64
1	8.6	23.04
1.5	7.8	13.89
2	12.4	16.57
2.5	14.4	15.39
3	18.3	16.3
3.5	15.1	11.51
4	24.7	16.51
4.5	15.6	9.24
5	14.8	7.89
6	16.6	7.37
7	15.3	5.82
8	11.9	3.96

Table VII. Frequency differences corresponding to threshold negative phase shifts.

	Ripple Density (ripples/octave)	Average Frequency DL	Threshold Phase shift
Flanagan study	0.5	55 Hz	10.1 degrees
	1.0	55 Hz	20.3 degrees
Ripple study	0.5	54 Hz	10.4 degrees
	1.0	21 Hz	7.8 degrees

Table VIII. Comparison of Flanagan DL values and ripple study threshold phase shift values.